

DEX-0214

EI

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
19 October 2000 (19.10.2000)

PCT

(10) International Publication Number
WO 00/61756 A3

(51) International Patent Classification⁷: C12N 15/12,
15/62, C07K 14/47, 16/18, A61K 35/12, 38/17, 39/00,
A61P 35/00, G01N 33/53, C12Q 1/68, A61K 48/00

(21) International Application Number: PCT/US00/09688

(22) International Filing Date: 10 April 2000 (10.04.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/288,950 9 April 1999 (09.04.1999) US
09/346,327 2 July 1999 (02.07.1999) US

(71) Applicant (for all designated States except US): CORIXA
CORPORATION [US/US]; Suite 200, 1124 Columbia
Street, Seattle, WA 98104 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): REED, Steven, G.
[US/US]; 2843 - 122nd Place NE, Bellevue, WA 98005
(US). XU, Jiangchun [US/US]; 15805 SE 43rd Place,
Bellevue, WA 98006 (US). DILLON, Davin, C. [US/US];
21607 NE 24th Street, Redmond, WA 98053 (US).

(74) Agents: POTTER, Jane, E.R.; Seed Intellectual Property
Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle,
WA 98104-7092 et al. (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE,
DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU,
ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ,
PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT,
TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR,
IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

(88) Date of publication of the international search report:
8 February 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 00/61756 A3

(54) Title: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE

(57) Abstract: Compounds and methods for the treatment and diagnosis of breast cancer are provided. The inventive compounds include polypeptides containing at least a portion of a breast tumor protein. Vaccines and pharmaceutical compositions for immunotherapy of breast cancer comprising such polypeptides, or polynucleotides encoding such polypeptides, are also provided, together with polynucleotides for preparing the inventive polypeptides.

INTERNATIONAL SEARCH REPORT

Intern. Appl. No
PCT/US 00/09688

A. CLASSIFICATION OF SUBJECT MATTER					
IPC 7	C12N15/12	C12N15/62	C07K14/47	C07K16/18	A61K35/12
	A61K38/17	A61K39/00	A61P35/00	G01N33/53	C12Q1/68
	A61K48/00				

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A61K A61P G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 679 716 A (MATSUBARA KENICHI ;OKUBO KOUSAKU (JP)) 2 November 1995 (1995-11-02) HUMGS01423/SeqIdNo.1275: 82.6% identity in 149 bp overlap with SeqIdNo.3 ---	1,3-12, 18-35, 37-60
A	DATABASE EMEST [Online] E.B.I., Hinxton, U.K.; Accession Number: AA749298, 20 January 1998 (1998-01-20) NCI-CGAP: "Homo sapiens cDNA clone IMAGE:1271152 3', mRNA sequence" XP002146686 abstract ---	
A	WO 95 19783 A (CELIS ESTEBAN ;GREY HOWARD M (US); CYTEL CORP (US); KUBO RALPH T () 27 July 1995 (1995-07-27) ---	

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

° Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

6 September 2000

Date of mailing of the international search report

04.12.2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Lonnay, O

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/09688

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SCHLOM J ET AL: "STRATEGIES FOR THE DEVELOPMENT OF RECOMBINANT VACCINES FOR THE IMMUNOTHERAPY OF BREAST CANCER" BREAST CANCER RESEARCH AND TREATMENT, US, NIJHOFF, BOSTON, vol. 38, no. 1, 1996, pages 27-39, XP000578043 ISSN: 0167-6806 ---	
A	WO 97 02280 A (HUMAN GENOME SCIENCES INC) 23 January 1997 (1997-01-23) ---	
P,X	WO 99 33869 A (CORIXA CORP) 8 July 1999 (1999-07-08) SeqIdNo.3: 100.0% identity in 421 bp overlap with SeqIdNo.3 -----	1,3-12, 18-35, 37-60

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/09688

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 48-53 and 56-58 are directed to methods of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1,3-12,18-35,37-60 (all partially)

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: Claims 1,3-12,18-35,37-60 (all partially)
An isolated polypeptide comprising an immunogenic portion of a breast protein or a variant thereof, encoded by a polynucleotide comprising the sequence of SeqIdNo.3; said polynucleotide; a vector and a host comprising the polynucleotide; its use for the production of the polypeptide or antibodies reactive with the polypeptide; its use for the manufacture of a pharmaceutical composition or a vaccine; its use for the detection of breast proteins or nucleic acids.

Inventions 2 to 104: Claims 1-60 (all partially, as applicable)
Idem as invention 1, wherein

Inventions 2-38 relate to the individual nucleic acids and their corresponding polypeptides in the order as they occur in claim 1, where invention 2 is limited to SeqIdNo.10, invention 3 to SeqIdNo.17, invention 4 to SeqIdNo.24, inventions 5-12 respectively to SeqIdNos.45-52, inventions 13-25 respectively to SeqIdNos.55-67, invention 26 to SeqIdNo.72, invention 27 to SeqIdNo.73, invention 28-33 respectively to SeqIdNos.89-94, invention 34 to SeqIdNo.95 and SeqIdNo.98, invention 35 to SeqIdNo.96 and 99, invention 36 to SeqIdNo.97, invention 37 to SeqIdNo.102, and invention 38 to SeqIdNo.107;

Inventions 39-103 relate to the individual nucleic acids and their corresponding polypeptides in the order as they occur in claim 13, where invention 39 is limited to SeqIdNo.1, invention 40 to SeqIdNo.2, inventions 41-46 to SeqIdNos.4-9, inventions 47-52 to SeqIdNo.11-16, inventions 53-58 to SeqIdNo.18-23, inventions 59-78 to SeqIdNo.25-44, invention 79 to SeqIdNo.53, invention 80 to SeqIdNo.54, inventions 81-84 to SeqIdNo.68-71, inventions 85-99 to SeqIdNo.74-88, and inventions 100-103 to SeqIdNo.103-106; and

Invention 104 relates to the individual nucleic acid of SeqIdNo.100 and the corresponding polypeptide of SeqIdNo.101.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/09688

Patent document cited in search report	Publication date	Patent family member(s)			Publication date
EP 0679716 A	02-11-1995	AU	8116494 A		13-06-1995
		CA	2153480 A		01-06-1995
		WO	9514772 A		01-06-1995

WO 9519783 A	27-07-1995	US	5662907 A		02-09-1997
		AU	1834095 A		08-08-1995
		AU	4111999 A		23-09-1999
		CA	2181920 A		27-07-1995
		EP	0749315 A		27-12-1996
		SG	49743 A		15-06-1998

WO 9702280 A	23-01-1997	CA	2225824 A		23-01-1997
		AU	3092995 A		05-02-1997
		EP	0851869 A		08-07-1998
		JP	11509093 T		17-08-1999

WO 9933869 A	08-07-1999	AU	2010699 A		19-07-1999
		EP	1042360 A		11-10-2000
		ZA	9811800 A		23-06-1999

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, 15/62, C07K 14/47, 16/18, A61K 35/12, 38/17, 39/00, A61P 35/00, G01N 33/53, C12Q 1/68, A61K 48/00		A2	(11) International Publication Number: WO 00/61756 (43) International Publication Date: 19 October 2000 (19.10.00)
 (21) International Application Number: PCT/US00/09688 (22) International Filing Date: 10 April 2000 (10.04.00) (30) Priority Data: 09/288,950 9 April 1999 (09.04.99) US 09/346,327 2 July 1999 (02.07.99) US		 (81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
 (71) Applicant (<i>for all designated States except US</i>): CORIXA CORPORATION [US/US]; Suite 200, 1124 Columbia Street, Seattle, WA 98104 (US). (72) Inventors; and (75) Inventors/Applicants (<i>for US only</i>): REED, Steven, G. [US/US]; 2843 - 122nd Place NE, Bellevue, WA 98005 (US). XU, Jiangchun [US/US]; 15805 SE 43rd Place, Bellevue, WA 98006 (US). DILLON, Davin, C. [US/US]; 21607 NE 24th Street, Redmond, WA 98053 (US). (74) Agents: POTTER, Jane, E.R.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 (US) et al.		 Published <i>Without international search report and to be republished upon receipt of that report.</i>	
 (54) Title: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE (57) Abstract Compounds and methods for the treatment and diagnosis of breast cancer are provided. The inventive compounds include polypeptides containing at least a portion of a breast tumor protein. Vaccines and pharmaceutical compositions for immunotherapy of breast cancer comprising such polypeptides, or polynucleotides encoding such polypeptides, are also provided, together with polynucleotides for preparing the inventive polypeptides.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
OF BREAST CANCER AND METHODS FOR THEIR USE

5 TECHNICAL FIELD

The present invention relates generally to compositions and methods for the treatment and diagnosis of breast cancer. The invention is more particularly related to polypeptides comprising at least a portion of a protein that is preferentially expressed in breast tumor tissue and to polynucleotides encoding such polypeptides.

10 Such polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for treatment of breast cancer. Additionally such polypeptides and polynucleotides may be used in the immunodiagnosis of breast cancer.

BACKGROUND OF THE INVENTION

15 Breast cancer is a significant health problem for women in the United States and throughout the world. Although advances have been made in detection and treatment of the disease, breast cancer remains the second leading cause of cancer-related deaths in women, affecting more than 180,000 women in the United States each year. For women in North America, the life-time odds of getting breast

20 cancer are now one in eight.

No vaccine or other universally successful method for the prevention or treatment of breast cancer is currently available. Management of the disease currently relies on a combination of early diagnosis (through routine breast screening procedures) and aggressive treatment, which may include one or more of a variety of

25 treatments such as surgery, radiotherapy, chemotherapy and hormone therapy. The course of treatment for a particular breast cancer is often selected based on a variety of prognostic parameters, including an analysis of specific tumor markers. See, e.g., Porter-Jordan and Lippman, *Breast Cancer* 8:73-100 (1994). However, the use of established markers often leads to a result that is difficult to interpret, and the high

mortality observed in breast cancer patients indicates that improvements are needed in the treatment, diagnosis and prevention of the disease.

Accordingly, there is a need in the art for improved methods for therapy and diagnosis of breast cancer. The present invention fulfills these needs and
5 further provides other related advantages.

SUMMARY OF THE INVENTION

The present invention provides compounds and methods for immunotherapy of breast cancer. In one aspect, isolated polypeptides are provided
10 comprising at least an immunogenic portion of a breast tumor protein or a variant of said protein that differs only in conservative substitutions and/or modifications, wherein the breast tumor protein comprises an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of (a) nucleotide sequences recited in SEQ ID NOS: 3, 10, 17, 24, 45-52, 55-67, 72, 73, 89-
15 97, 102 and 107, (b) complements of said nucleotide sequences and (c) sequences that hybridize to a sequence of (a) or (b) under moderately stringent conditions. In specific embodiments, the isolated polypeptides of the present invention comprise an amino acid sequence of SEQ ID NO: 98, 99 or 101.

In related aspects, isolated polynucleotides encoding the above
20 polypeptides are provided. In specific embodiments, such polynucleotides comprise sequences provided in SEQ ID NOS: 3, 10, 17, 24, 45-52 and 55-67, 72, 73, 89-97, 102 and 107. The present invention further provides expression vectors comprising the above polynucleotides and host cells transformed or transfected with such expression vectors. In preferred embodiments, the host cells are selected from the
25 group consisting of *E. coli*, yeast and mammalian cells.

In another aspect, the present invention provides fusion proteins comprising a first and a second inventive polypeptide or, alternatively, an inventive polypeptide and a known breast antigen.

The present invention also provides pharmaceutical compositions
30 comprising at least one of the above polypeptides, or a polynucleotide encoding such a polypeptide, and a physiologically acceptable carrier, together with vaccines

comprising at least one or more such polypeptide or polynucleotide in combination with a non-specific immune response enhancer. Pharmaceutical compositions and vaccines comprising one or more of the above fusion proteins are also provided.

In related aspects, pharmaceutical compositions for the treatment of
5 breast cancer comprising at least one polypeptide and a physiologically acceptable carrier are provided, wherein the polypeptide comprises an immunogenic portion of a breast tumor protein or a variant thereof, the breast tumor protein being encoded by a polynucleotide comprising a sequence selected from the group consisting of: (a) nucleotide sequences recited in SEQ ID NOS: 1, 2, 4-9, 11-16, 18-23, 25-44, 53, 54,
10 68-71, 74-88 and 103-106, (b) complements of said nucleotide sequences, and (c) sequences that hybridize to a sequence of (a) or (b) under moderately stringent conditions. The invention also provides vaccines for the treatment of breast cancer comprising such polypeptides in combination with a non-specific immune response enhancer, together with pharmaceutical compositions and vaccines comprising at
15 least one polynucleotide comprising a sequence provided in SEQ ID NOS: 1, 2, 4-9, 11-16, 18-23, 25-44, 53, 54, 68-71, 74-88 and 103-106.

In yet another aspect, methods are provided for inhibiting the development of breast cancer in a patient, comprising administering an effective amount of at least one of the above pharmaceutical compositions and/or vaccines.

20 The present invention also provides methods for immunodiagnosis of breast cancer, together with kits for use in such methods. In one specific aspect of the present invention, methods are provided for detecting breast cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that is capable of binding to one of the above polypeptides; and (b) detecting in
25 the sample a protein or polypeptide that binds to the binding agent. In preferred embodiments, the binding agent is an antibody, most preferably a monoclonal antibody.

In related aspects, methods are provided for monitoring the progression of breast cancer in a patient, comprising: (a) contacting a biological
30 sample obtained from a patient with a binding agent that is capable of binding to one

of the above polypeptides; (b) determining in the sample an amount of a protein or polypeptide that binds to the binding agent; (c) repeating steps (a) and (b); and comparing the amounts of polypeptide detected in steps (b) and (c).

Within related aspects, the present invention provides antibodies,
5 preferably monoclonal antibodies, that bind to the inventive polypeptides, as well as diagnostic kits comprising such antibodies, and methods of using such antibodies to inhibit the development of breast cancer.

The present invention further provides methods for detecting breast cancer comprising: (a) obtaining a biological sample from a patient; (b) contacting
10 the sample with a first and a second oligonucleotide primer in a polymerase chain reaction, at least one of the oligonucleotide primers being specific for a polynucleotide that encodes one of the above polypeptides; and (c) detecting in the sample a DNA sequence that amplifies in the presence of the first and second oligonucleotide primers. In a preferred embodiment, at least one of the
15 oligonucleotide primers comprises at least about 10 contiguous nucleotides of a polynucleotide comprising a sequence selected from the group consisting of SEQ ID NOS: 1-97, 100 and 102-107.

In a further aspect, the present invention provides a method for detecting breast cancer in a patient comprising: (a) obtaining a biological sample
20 from the patient; (b) contacting the sample with an oligonucleotide probe specific for a polynucleotide that encodes one of the above polypeptides; and (c) detecting in the sample a polynucleotide sequence that hybridizes to the oligonucleotide probe. Preferably, the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a polynucleotide comprising a sequence selected from the group
25 consisting of SEQ ID NOS: 1-97, 100 and 102-107.

In related aspects, diagnostic kits comprising the above oligonucleotide probes or primers are provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description. All references disclosed herein

are hereby incorporated by reference in their entirety as if each was incorporated individually.

5 BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE IDENTIFIERS

Figs. 1A and B show the specific lytic activity of a first and a second B511S-specific CTL clone, respectively, measured on autologous LCL transduced with B511s (filled squares) or HLA-A3 (open squares).

- 10 SEQ ID NO: 1 is the determined 3'cDNA sequence of 1T-5120
- SEQ ID NO: 2 is the determined 3'cDNA sequence of 1T-5122
- SEQ ID NO: 3 is the determined 3'cDNA sequence of 1T-5123
- SEQ ID NO: 4 is the determined 3'cDNA sequence of 1T-5125
- SEQ ID NO: 5 is the determined 3'cDNA sequence of 1T-5126
- 15 SEQ ID NO: 6 is the determined 3'cDNA sequence of 1T-5127
- SEQ ID NO: 7 is the determined 3'cDNA sequence of 1T-5129
- SEQ ID NO: 8 is the determined 3'cDNA sequence of 1T-5130
- SEQ ID NO: 9 is the determined 3'cDNA sequence of 1T-5133
- SEQ ID NO: 10 is the determined 3'cDNA sequence of 1T-5136
- 20 SEQ ID NO: 11 is the determined 3'cDNA sequence of 1T-5137
- SEQ ID NO: 12 is the determined 3'cDNA sequence of 1T-5139
- SEQ ID NO: 13 is the determined 3'cDNA sequence of 1T-5142
- SEQ ID NO: 14 is the determined 3'cDNA sequence of 1T-5143
- SEQ ID NO: 15 is the determined 5'cDNA sequence of 1T-5120
- 25 SEQ ID NO: 16 is the determined 5'cDNA sequence of 1T-5122
- SEQ ID NO: 17 is the determined 5'cDNA sequence of 1T-5123
- SEQ ID NO: 18 is the determined 5'cDNA sequence of 1T-5125
- SEQ ID NO: 19 is the determined 5'cDNA sequence of 1T-5126
- SEQ ID NO: 20 is the determined 5'cDNA sequence of 1T-5127
- 30 SEQ ID NO: 21 is the determined 5'cDNA sequence of 1T-5129
- SEQ ID NO: 22 is the determined 5'cDNA sequence of 1T-5130

- SEQ ID NO: 23 is the determined 5'cDNA sequence of 1T-5133
SEQ ID NO: 24 is the determined 5'cDNA sequence of 1T-5136
SEQ ID NO: 25 is the determined 5'cDNA sequence of 1T-5137
SEQ ID NO: 26 is the determined 5'cDNA sequence of 1T-5139
5 SEQ ID NO: 27 is the determined 5'cDNA sequence of 1T-5142
SEQ ID NO: 28 is the determined 5'cDNA sequence of 1T-5143
SEQ ID NO: 29 is the determined 5'cDNA sequence of 1D-4315
SEQ ID NO: 30 is the determined 5'cDNA sequence of 1D-4311
SEQ ID NO: 31 is the determined 5'cDNA sequence of 1E-4440
10 SEQ ID NO: 32 is the determined 5'cDNA sequence of 1E-4443
SEQ ID NO: 33 is the determined 5'cDNA sequence of 1D-4321
SEQ ID NO: 34 is the determined 5'cDNA sequence of 1D-4310
SEQ ID NO: 35 is the determined 5'cDNA sequence of 1D-4320
SEQ ID NO: 36 is the determined 5'cDNA sequence of 1E-4448
15 SEQ ID NO: 37 is the determined 5'cDNA sequence of 1S-5105
SEQ ID NO: 38 is the determined 5'cDNA sequence of 1S-5110
SEQ ID NO: 39 is the determined 5'cDNA sequence of 1S-5111
SEQ ID NO: 40 is the determined 5'cDNA sequence of 1S-5116
SEQ ID NO: 41 is the determined 5'cDNA sequence of 1S-5114
20 SEQ ID NO: 42 is the determined 5'cDNA sequence of 1S-5115
SEQ ID NO: 43 is the determined 5'cDNA sequence of 1S-5118
SEQ ID NO: 44 is the determined 5'cDNA sequence of 1T-5134
SEQ ID NO: 45 is the determined 5'cDNA sequence of 1E-4441
SEQ ID NO: 46 is the determined 5'cDNA sequence of 1E-4444
25 SEQ ID NO: 47 is the determined 5'cDNA sequence of 1E-4322
SEQ ID NO: 48 is the determined 5'cDNA sequence of 1S-5103
SEQ ID NO: 49 is the determined 5'cDNA sequence of 1S-5107
SEQ ID NO: 50 is the determined 5'cDNA sequence of 1S-5113
SEQ ID NO: 51 is the determined 5'cDNA sequence of 1S-5117
30 SEQ ID NO: 52 is the determined 5'cDNA sequence of 1S-5112

- SEQ ID NO: 53 is the determined cDNA sequence of 1013E11
SEQ ID NO: 54 is the determined cDNA sequence of 1013H10
SEQ ID NO: 55 is the determined cDNA sequence of 1017C2
SEQ ID NO: 56 is the determined cDNA sequence of 1016F8
5 SEQ ID NO: 57 is the determined cDNA sequence of 1015F5
SEQ ID NO: 58 is the determined cDNA sequence of 1017A11
SEQ ID NO: 59 is the determined cDNA sequence of 1013A11
SEQ ID NO: 60 is the determined cDNA sequence of 1016D8
SEQ ID NO: 61 is the determined cDNA sequence of 1016D12
10 SEQ ID NO: 62 is the determined cDNA sequence of 1015E8
SEQ ID NO: 63 is the determined cDNA sequence of 1015D11
SEQ ID NO: 64 is the determined cDNA sequence of 1012H8
SEQ ID NO: 65 is the determined cDNA sequence of 1013C8
SEQ ID NO: 66 is the determined cDNA sequence of 1014B3
15 SEQ ID NO: 67 is the determined cDNA sequence of 1015B2
SEQ ID NO: 68-71 are the determined cDNA sequences of previously identified
antigens
SEQ ID NO: 72 is the determined cDNA sequence of JJ9434
SEQ ID NO: 73 is the determined cDNA sequence of B535S
20 SEQ ID NO: 74-88 are the determined cDNA sequence of previously identified
antigens
SEQ ID NO: 89 is the determined cDNA sequence of B534S
SEQ ID NO: 90 is the determined cDNA sequence of B538S
SEQ ID NO: 91 is the determined cDNA sequence of B542S
25 SEQ ID NO: 92 is the determined cDNA sequence of B543S
SEQ ID NO: 93 is the determined cDNA sequence of P501S
SEQ ID NO: 94 is the determined cDNA sequence of B541S
SEQ ID NO: 95 is an extended cDNA sequence for 1016F8 (also referred to as
B511S)
30 SEQ ID NO: 96 is an extended cDNA sequence for 1016D12 (also referred to as

B532S)

SEQ ID NO: 97 is an extended cDNA sequence for 1012H8 (also referred to as B533S)

SEQ ID NO: 98 is the predicted amino acid sequence for B511S

5 SEQ ID NO: 99 is the predicted amino acid sequence for B532S

SEQ ID NO: 100 is the determined full-length cDNA sequence for P501S

SEQ ID NO: 101 is the predicted amino acid sequence for P501S

SEQ ID NO: 102 is the determined cDNA sequence of clone #19605, also referred to as 1017C2, showing no significant homology to any known gene

10 SEQ ID NO: 103 is the determined 3' end cDNA sequence for clone #19599, showing homology to the Tumor Expression Enhanced gene

SEQ ID NO: 104 is the determined 5' end cDNA sequence for clone #19599, showing homology to the Tumor Expression Enhanced gene

SEQ ID NO: 105 is the determined cDNA sequence for clone #19607, showing

15 homology to Stromelysin-3

SEQ ID NO: 106 is the determined cDNA sequence for clone #19601, showing homology to Collagen

SEQ ID NO: 107 is the determined cDNA sequence of clone #19606, also referred to as B546S, showing no significant homology to any known gene

20

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for the immunotherapy and diagnosis of breast cancer.

The inventive compositions are generally isolated polypeptides that comprise at least

25 a portion of a breast tumor protein. Also included within the present invention are molecules (such as an antibody or fragment thereof) that bind to the inventive polypeptides. Such molecules are referred to herein as "binding agents."

In particular, the subject invention discloses isolated polypeptides comprising at least a portion of a human breast tumor protein, or a variant thereof,

30 wherein the breast tumor protein includes an amino acid sequence encoded by a

polynucleotide molecule including a sequence selected from the group consisting of: nucleotide sequences recited in SEQ ID NOS: 1-97, 100 and 102-107, the complements of said nucleotide sequences, and variants thereof. In certain specific embodiments, the inventive polypeptides comprise an amino acid sequence selected
5 from the group consisting of sequences provided in SEQ ID NO: 98, 99 and 101, and variants thereof. As used herein, the term "polypeptide" encompasses amino acid chains of any length, including full length proteins, wherein the amino acid residues are linked by covalent peptide bonds. Thus, a polypeptide comprising a portion of one of the above breast proteins may consist entirely of the portion, or the portion
10 may be present within a larger polypeptide that contains additional sequences. The additional sequences may be derived from the native protein or may be heterologous, and such sequences may be immunoreactive and/or antigenic.

As used herein, an "immunogenic portion" of a human breast tumor protein is a portion that is capable of eliciting an immune response in a patient inflicted with breast cancer and as such binds to antibodies present within sera from a breast cancer patient. Such immunogenic portions generally comprise at least about 5 amino acid residues, more preferably at least about 10, and most preferably at least about 20 amino acid residues. Immunogenic portions of the proteins described herein may be identified in antibody binding assays. Such assays may generally be
15 performed using any of a variety of means known to those of ordinary skill in the art, as described, for example, in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988. For example, a polypeptide may be immobilized on a solid support (as described below) and contacted with patient sera to allow binding of antibodies within the sera to the
20 immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ¹²⁵I-labeled Protein A. Alternatively, a polypeptide may be used to generate monoclonal and polyclonal antibodies for use in detection of the polypeptide in blood or other fluids of breast cancer patients. Methods for preparing and identifying immunogenic portions of antigens of known sequence are well known
25

in the art and include those summarized in Paul, *Fundamental Immunology*, 3rd ed., Raven Press, 1993, pp. 243-247.

The term "polynucleotide(s)," as used herein, means a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases and includes 5 DNA and corresponding RNA molecules, including HnRNA and mRNA molecules, both sense and anti-sense strands, and comprehends cDNA, genomic DNA and recombinant DNA, as well as wholly or partially synthesized polynucleotides. An HnRNA molecule contains introns and corresponds to a DNA molecule in a generally one-to-one manner. An mRNA molecule corresponds to an HnRNA and DNA 10 molecule from which the introns have been excised. A polynucleotide may consist of an entire gene, or any portion thereof. Operable anti-sense polynucleotides may comprise a fragment of the corresponding polynucleotide, and the definition of "polynucleotide" therefore includes all such operable anti-sense fragments.

The compositions and methods of the present invention also 15 encompass variants of the above polypeptides and polynucleotides. A polypeptide "variant," as used herein, is a polypeptide that differs from the recited polypeptide only in conservative substitutions and/or modifications, such that the therapeutic, antigenic and/or immunogenic properties of the polypeptide are retained. In a preferred embodiment, variant polypeptides differ from an identified sequence by 20 substitution, deletion or addition of five amino acids or fewer. Such variants may generally be identified by modifying one of the above polypeptide sequences, and evaluating the antigenic properties of the modified polypeptide using, for example, the representative procedures described herein. Polypeptide variants preferably exhibit at least about 70%, more preferably at least about 90% and most preferably at 25 least about 95% identity (determined as described below) to the identified polypeptides.

As used herein, a "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and 30 hydrophobic nature of the polypeptide to be substantially unchanged. In general, the

11

following groups of amino acids represent conservative changes: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his.

Variants may also, or alternatively, contain other modifications, 5 including the deletion or addition of amino acids that have minimal influence on the antigenic properties, secondary structure and hydropathic nature of the polypeptide. For example, a polypeptide may be conjugated to a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other 10 sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

A nucleotide "variant" is a sequence that differs from the recited 15 nucleotide sequence in having one or more nucleotide deletions, substitutions or additions. Such modifications may be readily introduced using standard mutagenesis techniques, such as oligonucleotide-directed site-specific mutagenesis as taught, for example, by Adelman et al. (*DNA*, 2:183, 1983). Nucleotide variants may be naturally occurring allelic variants, or non-naturally occurring variants. Variant nucleotide sequences preferably exhibit at least about 70%, more preferably at least 20 about 80% and most preferably at least about 90% identity (determined as described below) to the recited sequence.

The antigens provided by the present invention include variants that are encoded by DNA sequences which are substantially homologous to one or more of the DNA sequences specifically recited herein. "Substantial homology," as used 25 herein, refers to DNA sequences that are capable of hybridizing under moderately stringent conditions. Suitable moderately stringent conditions include prewashing in a solution of 5X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5X SSC, overnight or, in the event of cross-species homology, at 45°C with 0.5X SSC; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 30 0.2X SSC containing 0.1% SDS. Such hybridizing DNA sequences are also within

the scope of this invention, as are nucleotide sequences that, due to code degeneracy, encode an immunogenic polypeptide that is encoded by a hybridizing DNA sequence.

Two nucleotide or polypeptide sequences are said to be "identical" if the sequence of nucleotides or amino acid residues in the two sequences is the same 5 when aligned for maximum correspondence as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, more preferably 40 to about 50, in 10 which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies 15 several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) *Unified Approach to Alignment and Phylogenies* pp. 626-645 20 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) Fast and sensitive multiple sequence alignments on a microcomputer *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) Optimal alignments in linear space *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Saitou, N. Nes, M. (1987) The neighbor joining method. A new method for 25 reconstructing phylogenetic trees *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) Rapid similarity searches of nucleic acid and protein data banks *Proc. Natl. Acad. Sci. USA* 80:726-730.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e. gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the 10 total number of positions in the reference sequence (i.e. the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Also included in the scope of the present invention are alleles of the genes encoding the nucleotide sequences recited herein. As used herein, an "allele" or "allelic sequence" is an alternative form of the gene which may result from at least 15 one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one, or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone or in combination with 20 the others, one or more times in a given sequence.

For breast tumor polypeptides with immunoreactive properties, variants may alternatively be identified by modifying the amino acid sequence of one of the above polypeptides, and evaluating the immunoreactivity of the modified polypeptide. For breast tumor polypeptides useful for the generation of diagnostic binding agents, a 25 variant may be identified by evaluating a modified polypeptide for the ability to generate antibodies that detect the presence or absence of breast cancer. Such modified sequences may be prepared and tested using, for example, the representative procedures described herein.

The breast tumor proteins of the present invention, and polynucleotide 30 molecules encoding such proteins, may be isolated from breast tumor tissue using any

of a variety of methods well known in the art. Polynucleotide sequences corresponding to a gene (or a portion thereof) encoding one of the inventive breast tumor proteins may be isolated from a breast tumor cDNA library using a subtraction technique as described in detail below. Examples of such DNA sequences are 5 provided in SEQ ID NOS: 1- 97, 100 and 102-107. Partial polynucleotide sequences thus obtained may be used to design oligonucleotide primers for the amplification of full-length polynucleotide sequences in a polymerase chain reaction (PCR), using techniques well known in the art (see, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 10 1989). Once a polynucleotide sequence encoding a polypeptide is obtained, any of the above modifications may be readily introduced using standard mutagenesis techniques, such as oligonucleotide-directed site-specific mutagenesis as taught, for example, by Adelman et al. (*DNA*, 2:183, 1983).

The breast tumor polypeptides disclosed herein may also be generated 15 by synthetic or recombinant means. Synthetic polypeptides having fewer than about 100 amino acids, and generally fewer than about 50 amino acids, may be generated using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids 20 are sequentially added to a growing amino acid chain (see, for example, Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963). Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

25 Alternatively, any of the above polypeptides may be produced recombinantly by inserting a polynucleotide sequence that encodes the polypeptide into an expression vector and expressing the protein in an appropriate host. Any of a variety of expression vectors known to those of ordinary skill in the art may be employed to express recombinant polypeptides of this invention. Expression may be 30 achieved in any appropriate host cell that has been transformed or transfected with an

expression vector containing a polynucleotide molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast and higher eukaryotic cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line, such as CHO cells. The polynucleotide sequences expressed in this manner may 5 encode naturally occurring polypeptides, portions of naturally occurring polypeptides, or other variants thereof.

In general, regardless of the method of preparation, the polypeptides disclosed herein are prepared in an isolated, substantially pure form (*i.e.*, the polypeptides are homogenous as determined by amino acid composition and primary 10 sequence analysis). Preferably, the polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. In certain preferred embodiments, described in more detail below, the substantially pure polypeptides are incorporated into pharmaceutical compositions or vaccines for use in one or more of the methods disclosed herein.

15 In a related aspect, the present invention provides fusion proteins comprising a first and a second inventive polypeptide or, alternatively, a polypeptide of the present invention and a known breast tumor antigen, together with variants of such fusion proteins.

A polynucleotide sequence encoding a fusion protein of the present 20 invention is constructed using known recombinant DNA techniques to assemble separate polynucleotide sequences encoding the first and second polypeptides into an appropriate expression vector. The 3' end of a polynucleotide sequence encoding the first polypeptide is ligated, with or without a peptide linker, to the 5' end of a polynucleotide sequence encoding the second polypeptide so that the reading frames 25 of the sequences are in phase to permit mRNA translation of the two DNA sequences into a single fusion protein that retains the biological activity of both the first and the second polypeptides.

A peptide linker sequence may be employed to separate the first and the second polypeptides by a distance sufficient to ensure that each polypeptide folds 30 into its secondary and tertiary structures. Such a peptide linker sequence is

incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and 5 second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 10 Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may be from 1 to about 50 amino acids in length. Peptide sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

15 The ligated polynucleotide sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of polynucleotides are located only 5' to the polynucleotide sequence encoding the first polypeptide. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the 20 polynucleotide sequence encoding the second polypeptide.

Fusion proteins are also provided that comprise a polypeptide of the present invention together with an unrelated immunogenic protein. Preferably the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (see, for example, Stoute 25 et al. *New Engl. J. Med.*, 336:86-91 (1997)).

Polypeptides of the present invention that comprise an immunogenic portion of a breast tumor protein may generally be used for immunotherapy of breast cancer, wherein the polypeptide stimulates the patient's own immune response to breast tumor cells. In further aspects, the present invention provides methods for 30 using one or more of the immunoreactive polypeptides encoded by a polynucleotide

molecule having a sequence provided in SEQ ID NOS: 1- 97, 100 and 102-107 (or fusion proteins comprising one or more such polypeptides and/or polynucleotides encoding such polypeptides) for immunotherapy of breast cancer in a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. A 5 patient may be afflicted with a disease, or may be free of detectable disease. Accordingly, the above immunoreactive polypeptides (or fusion proteins or polynucleotide molecules encoding such polypeptides) may be used to treat breast cancer or to inhibit the development of breast cancer. In a preferred embodiment, the polypeptides are administered either prior to or following surgical removal of primary 10 tumors and/or treatment by administration of radiotherapy and conventional chemotherapeutic drugs.

In these aspects, the polypeptide or fusion protein is generally present within a pharmaceutical composition and/or a vaccine. Pharmaceutical compositions may comprise one or more polypeptides, each of which may contain one or more of 15 the above sequences (or variants thereof), and a physiologically acceptable carrier. The vaccines may comprise one or more of such polypeptides and a non-specific immune response enhancer, wherein the non-specific immune response enhancer is capable of eliciting or enhancing an immune response to an exogenous antigen. Examples of non-specific-immune response enhancers include adjuvants, 20 biodegradable microspheres (*e.g.*, polylactic galactide) and liposomes (into which the polypeptide is incorporated). Pharmaceutical compositions and vaccines may also contain other epitopes of breast tumor antigens, either incorporated into a combination polypeptide (*i.e.*, a single polypeptide that contains multiple epitopes) or present within a separate polypeptide.

25 Alternatively, a pharmaceutical composition or vaccine may contain polynucleotides encoding one or more of the above polypeptides, such that the polypeptide is generated *in situ*. In such pharmaceutical compositions and vaccines, the polynucleotide may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, 30 bacteria and viral expression systems. Appropriate nucleic acid expression systems

contain the necessary polynucleotide sequences for expression in the patient (such as a suitable promoter). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an epitope of a breast tumor cell antigen on its cell surface. In a preferred embodiment, the polynucleotide molecules may be introduced using a viral expression system (e.g., vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch et al., *PNAS* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 10 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991; Kolls et al., *PNAS* 91:215-219, 1994; Kass-Eisler et al., *PNAS* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993. Techniques for incorporating 15 polynucleotides into such expression systems are well known to those of ordinary skill in the art.

The polynucleotides may also be "naked," as described, for example, in published PCT application WO 90/11092, and Ulmer et al., *Science* 259:1745-1749, 1993, reviewed by Cohen, *Science* 259:1691-1692, 1993. The 20 uptake of naked polynucleotides may be increased by coating the polynucleotides onto biodegradable beads, which are efficiently transported into the cells.

Routes and frequency of administration, as well as dosage, will vary from individual to individual and may parallel those currently being used in immunotherapy of other diseases. In general, the pharmaceutical compositions and 25 vaccines may be administered by injection (e.g., intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (e.g., by aspiration) or orally. Between 1 and 10 doses may be administered over a 3-24 week period. Preferably, 4 doses are administered, at an interval of 3 months, and booster administrations may be given periodically thereafter. Alternate protocols may be appropriate for individual 30 patients. A suitable dose is an amount of polypeptide or polynucleotide that is

effective to raise an immune response (cellular and/or humoral) against breast tumor cells in a treated patient. A suitable immune response is at least 10-50% above the basal (*i.e.*, untreated) level. In general, the amount of polypeptide present in a dose (or produced *in situ* by the polynucleotide in a dose) ranges from about 1 pg to about 5 100 mg per kg of host, typically from about 10 pg to about 1 mg, and preferably from about 100 pg to about 1 μ g. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.01 mL to about 5 mL.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of 10 carrier will vary depending on the mode of administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a lipid, a wax and/or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and/or magnesium 15 carbonate, may be employed. Biodegradable microspheres (*e.g.*, polylactic glycolide) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268 and 5,075,109.

Any of a variety of non-specific immune response enhancers may be 20 employed in the vaccines of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a nonspecific stimulator of immune response, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis*. Such adjuvants are commercially available as, for example, Freund's 25 Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI) and Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ).

Polypeptides disclosed herein may also be employed in adoptive immunotherapy for the treatment of cancer. Adoptive immunotherapy may be broadly classified into either active or passive immunotherapy. In active 30 immunotherapy, treatment relies on the *in vivo* stimulation of the endogenous host

immune system to react against tumors with the administration of immune response-modifying agents (for example, tumor vaccines, bacterial adjuvants, and/or cytokines).

In passive immunotherapy, treatment involves the delivery of biologic reagents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T lymphocytes (for example, CD8+ cytotoxic T-lymphocyte, CD4+ T-helper, gamma/delta T lymphocytes, tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells, lymphokine-activated killer cells), B cells, or antigen presenting cells (such as dendritic cells and macrophages) expressing the disclosed antigens. The polypeptides disclosed herein may also be used to generate antibodies or anti-idiotypic antibodies (as in U.S. Patent No. 4,918,164), for passive immunotherapy.

The predominant method of procuring adequate numbers of T-cells for adoptive immunotherapy is to grow immune T-cells *in vitro*. Culture conditions for expanding single antigen-specific T-cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. These *in vitro* culture conditions typically utilize intermittent stimulation with antigen, often in the presence of cytokines, such as IL-2, and non-dividing feeder cells. As noted above, the immunoreactive polypeptides described herein may be used to rapidly expand antigen-specific T cell cultures in order to generate sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast or B-cells, may be pulsed with immunoreactive polypeptides or polynucleotide sequence(s) may be introduced into antigen presenting cells, using standard techniques well known in the art. For example, antigen presenting cells may be transfected or transduced with a polynucleotide sequence, wherein said sequence contains a promoter region appropriate for inducing expression, and can be expressed as part of a recombinant virus or other expression system. Several viral vectors may be used to transduce an antigen presenting cell, including pox virus, vaccinia virus, and adenovirus. Antigen presenting cells may be

transfected with polynucleotide sequences disclosed herein by a variety of means, including gene-gun technology, lipid-mediated delivery, electroporation, osmotic shock, and particulate delivery mechanisms, resulting in efficient and acceptable expression levels as determined by one of ordinary skill in the art. For cultured T-
5 cells to be effective in therapy, the cultured T-cells must be able to grow and distribute widely and to survive long term *in vivo*. Studies have demonstrated that cultured T-cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (see, for example, Cheever, M., et al, "Therapy With Cultured T Cells: Principles Revisited,"

10 *Immunological Reviews*, 157:177, 1997).

The polypeptides disclosed herein may also be employed to generate and/or isolate tumor-reactive T-cells, which can then be administered to the patient. In one technique, antigen-specific T-cell lines may be generated by *in vivo* immunization with short peptides corresponding to immunogenic portions of the
15 disclosed polypeptides. The resulting antigen specific CD8+ CTL clones may be isolated from the patient, expanded using standard tissue culture techniques, and returned to the patient.

Alternatively, peptides corresponding to immunogenic portions of the polypeptides may be employed to generate tumor reactive T cell subsets by selective
20 *in vitro* stimulation and expansion of autologous T cells to provide antigen-specific T cells which may be subsequently transferred to the patient as described, for example, by Chang et al. (*Crit. Rev. Oncol. Hematol.*, 22(3), 213, 1996). Cells of the immune system, such as T cells, may be isolated from the peripheral blood of a patient, using a commercially available cell separation system. The separated cells are stimulated
25 with one or more of the immunoreactive polypeptides contained within a delivery vehicle, such as a microsphere, to provide antigen-specific T cells. The population of tumor antigen-specific T cells is then expanded using standard techniques and the cells are administered back to the patient.

In other embodiments, T-cell and/or antibody receptors specific for the
30 polypeptides can be cloned, expanded, and transferred into other vectors or effector

cells for use in adoptive immunotherapy. In particular, T cells may be transfected with the appropriate genes to express the variable domains from tumor specific monoclonal antibodies as the extracellular recognition elements and joined to the T cell receptor signaling chains, resulting in T cell activation, specific lysis, and 5 cytokine release. This enables the T cell to redirect its specificity in an MHC-independent manner. See for example, Eshhar, Z., *Cancer Immunol Immunother*, 45(3-4):131-6, 1997 and Hwu, P., et al, *Cancer Res*, 55(15):3369-73, 1995. Another embodiment may include the transfection of tumor antigen specific alpha and beta T cell receptor chains into alternate T cells, as in Cole, DJ, et al, *Cancer Res*, 55(4):748-10 52, 1995.

In further embodiments, syngeneic or autologous dendritic cells may be pulsed with peptides corresponding to at least an immunogenic portion of a polypeptide disclosed herein. The resulting antigen-specific dendritic cells may either be transferred into a patient, or employed to stimulate T cells to provide 15 antigen-specific T cells which may, in turn, be administered to a patient. The use of peptide-pulsed dendritic cells to generate antigen-specific T cells and the subsequent use of such antigen-specific T cells to eradicate tumors in a murine model has been demonstrated by Cheever et al. (*Immunological Reviews*, 157:177, 1997).

Additionally, vectors expressing the disclosed polynucleotides may be 20 introduced into stem cells taken from the patient and clonally propagated *in vitro* for autologous transplant back into the same patient.

In one specific embodiment, cells of the immune system, such as T cells, may be isolated from the peripheral blood of a patient, using a commercially available cell separation system, such as CellPro Incorporated's (Bothell, WA) 25 CEPRATE™ system (see U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). The separated cells are stimulated with one or more of the immunoreactive polypeptides contained within a delivery vehicle, such as a microsphere, to provide antigen-specific T cells. The population of tumor 30 antigen-specific T cells is then expanded using standard techniques and the cells are administered back to the patient.

Polypeptides of the present invention may also, or alternatively, be used to generate binding agents, such as antibodies or fragments thereof, that are capable of detecting metastatic human breast tumors. Binding agents of the present invention may generally be prepared using methods known to those of ordinary skill 5 in the art, including the representative procedures described herein. Binding agents are capable of differentiating between patients with and without breast cancer, using the representative assays described herein. In other words, antibodies or other binding agents raised against a breast tumor protein, or a suitable portion thereof, will generate a signal indicating the presence of primary or metastatic breast cancer in at 10 least about 20% of patients afflicted with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without primary or metastatic breast cancer. Suitable portions of such breast tumor proteins are portions that are able to generate a binding agent that indicates the presence of primary or metastatic breast cancer in substantially all (*i.e.*, at least about 15 80%, and preferably at least about 90%) of the patients for which breast cancer would be indicated using the full length protein, and that indicate the absence of breast cancer in substantially all of those samples that would be negative when tested with full length protein. The representative assays described below, such as the two-antibody sandwich assay, may generally be employed for evaluating the ability of a 20 binding agent to detect metastatic human breast tumors.

The ability of a polypeptide prepared as described herein to generate antibodies capable of detecting primary or metastatic human breast tumors may generally be evaluated by raising one or more antibodies against the polypeptide (using, for example, a representative method described herein) and determining the 25 ability of such antibodies to detect such tumors in patients. This determination may be made by assaying biological samples from patients with and without primary or metastatic breast cancer for the presence of a polypeptide that binds to the generated antibodies. Such test assays may be performed, for example, using a representative procedure described below. Polypeptides that generate antibodies capable of 30 detecting at least 20% of primary or metastatic breast tumors by such procedures are

considered to be useful in assays for detecting primary or metastatic human breast tumors. Polypeptide specific antibodies may be used alone or in combination to improve sensitivity.

Polypeptides capable of detecting primary or metastatic human breast
5 tumors may be used as markers for diagnosing breast cancer or for monitoring disease progression in patients. In one embodiment, breast cancer in a patient may be diagnosed by evaluating a biological sample obtained from the patient for the level of one or more of the above polypeptides, relative to a predetermined cut-off value. As used herein, suitable "biological samples" include blood, sera and urine.

10 The level of one or more of the above polypeptides may be evaluated using any binding agent specific for the polypeptide(s). A "binding agent," in the context of this invention, is any agent (such as a compound or a cell) that binds to a polypeptide as described above. As used herein, "binding" refers to a noncovalent association between two separate molecules (each of which may be free (*i.e.*, in solution) or present on the surface of a cell or a solid support), such that a "complex" 15 is formed. Such a complex may be free or immobilized (either covalently or noncovalently) on a support material. The ability to bind may generally be evaluated by determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by
20 the product of the component concentrations. In general, two compounds are said to "bind" in the context of the present invention when the binding constant for complex formation exceeds about 10^3 L/mol. The binding constant may be determined using methods well known to those of ordinary skill in the art.

Any agent that satisfies the above requirements may be a binding
25 agent. For example, a binding agent may be a ribosome with or without a peptide component, an RNA molecule or a peptide. In a preferred embodiment, the binding partner is an antibody, or a fragment thereof. Such antibodies may be polyclonal, or monoclonal. In addition, the antibodies may be single chain, chimeric, CDR-grafted or humanized. Antibodies may be prepared by the methods described herein and by
30 other methods well known to those of skill in the art.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding partner to detect polypeptide markers in a sample. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In a preferred embodiment, the assay involves the use of binding 5 partner immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a second binding partner that contains a reporter group. Suitable second binding partners include antibodies that bind to the binding partner/polypeptide complex. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled 10 with a reporter group and allowed to bind to the immobilized binding partner after incubation of the binding partner with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding partner is indicative of the reactivity of the sample with the immobilized binding partner.

The solid support may be any material known to those of ordinary skill 15 in the art to which the antigen may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. 20 Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the 25 antigen and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour 30 and about 1 day. In general, contacting a well of a plastic microtiter plate (such as

polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 µg, and preferably about 100 ng to about 1 µg, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally 5 be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active 10 hydrogen on the binding partner (see, e.g., Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the 15 sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a second antibody (containing a reporter group) capable of binding to a different site on the polypeptide is added. The amount of second antibody that remains bound to the solid support is then determined using a 20 method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20TM (Sigma Chemical Co., St. Louis, MO). The 25 immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is that period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with breast 30 cancer. Preferably, the contact time is sufficient to achieve a level of binding that is

at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about
5 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20TM. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include enzymes (such as horseradish peroxidase),
10 substrates, cofactors, inhibitors, dyes, radionuclides, luminescent groups, fluorescent groups and biotin. The conjugation of antibody to reporter group may be achieved using standard methods known to those of ordinary skill in the art.

The second antibody is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound second antibody is then removed and bound second antibody is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or
15 autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time),
20 followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of breast cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value is the average mean signal obtained when the
30 immobilized antibody is incubated with samples from patients without breast cancer.

In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for breast cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for breast cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the antibody is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized antibody as the sample passes through the membrane. A second, labeled antibody then binds to the antibody-polypeptide complex as a solution containing the second antibody flows through the membrane. The detection of bound second antibody may then be performed as described above. In the strip test format, one end of the membrane to which antibody is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second antibody and to the area of immobilized antibody. Concentration of second antibody at the area of immobilized antibody indicates the presence of breast cancer. Typically, the concentration of second antibody at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of antibody immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level

of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 μ g, and more preferably from about 50 ng to about 500 ng. Such tests can typically be
5 performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the antigens or antibodies of the present invention. The above descriptions are intended to be exemplary only.

In another embodiment, the above polypeptides may be used as
10 markers for the progression of breast cancer. In this embodiment, assays as described above for the diagnosis of breast cancer may be performed over time, and the change in the level of reactive polypeptide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, breast cancer is progressing in those patients in
15 whom the level of polypeptide detected by the binding agent increases over time. In contrast, breast cancer is not progressing when the level of reactive polypeptide either remains constant or decreases with time.

Antibodies for use in the above methods may be prepared by any of a variety of techniques known to those of ordinary skill in the art. See, e.g., Harlow
20 and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In one such technique, an immunogen comprising the antigenic polypeptide is initially injected into any of a wide variety of mammals (e.g., mice, rats, rabbits, sheep and goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short
25 polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may

then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for the antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein,
5 *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for
10 example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection
15 technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of
20 growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel
25 filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

Monoclonal antibodies of the present invention may also be used as therapeutic reagents, to diminish or eliminate breast tumors. The antibodies may be used on their own (for instance, to inhibit metastases) or coupled to one or more
30 therapeutic agents. Suitable agents in this regard include radionuclides,

differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ⁹⁰Y, ¹²³I, ¹²⁵I, ¹³¹I, ¹⁸⁶Re, ¹⁸⁸Re, ²¹¹At, and ²¹²Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, 5 diphtheria toxin, cholera toxin, gelonin, Pseudomonas exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a 10 substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulphhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an 15 antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on 20 agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino 25 groups, carboxyl groups, sulphhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a 30 linker group which is cleavable during or upon internalization into a cell. A number

of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitzer), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers which provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous,

intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

5 Diagnostic reagents of the present invention may also comprise at least a portion of a polynucleotide disclosed herein. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify breast tumor-specific cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for a polynucleotide
10 encoding a breast tumor protein of the present invention. The presence of the amplified cDNA is then detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes specific for a polynucleotide encoding a breast tumor protein of the present invention may be used in a hybridization assay to detect the presence of an inventive polypeptide in a
15 biological sample.

As used herein, the term "oligonucleotide primer/probe specific for a polynucleotide" means an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to the polynucleotide in question, or an oligonucleotide sequence that is anti-sense to a
20 sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to the polynucleotide in question. Oligonucleotide primers and/or probes which may be usefully employed in the inventive diagnostic methods preferably have at least about 10-40 nucleotides. In a preferred embodiment, the oligonucleotide primers comprise at least about 10
25 contiguous nucleotides of a polynucleotide disclosed herein or that is anti-sense to a polynucleotide sequence disclosed herein. Preferably, oligonucleotide probes for use in the inventive diagnostic methods comprise at least about 15 contiguous oligonucleotides of a polynucleotide that encodes one of the polypeptides disclosed herein or that is anti-sense to a sequence that encodes one of the polypeptides
30 disclosed herein. Techniques for both PCR based assays and hybridization assays are

well known in the art (see, for example, Mullis *et al. Ibid*; Ehrlich, *Ibid*). Primers or probes may thus be used to detect breast tumor-specific sequences in biological samples, including blood, urine and/or breast tumor tissue.

The following Examples are offered by way of illustration and not by
5 way of limitation.

EXAMPLES

Example 1

10 ISOLATION AND CHARACTERIZATION OF BREAST TUMOR POLYPEPTIDES

This Example describes the isolation of breast tumor polypeptides from a breast tumor cDNA library.

15 A human breast tumor cDNA expression library was constructed from a pool of breast tumor poly A⁺ RNA from three patients using a Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning kit (BRL Life Technologies, Gaithersburg, MD 20897) following the manufacturer's protocol. Specifically, breast tumor tissues were homogenized with polytron (Kinematica, Switzerland) and total
20 RNA was extracted using Trizol reagent (BRL Life Technologies) as directed by the manufacturer. The poly A⁺ RNA was then purified using a Qiagen oligotex spin column mRNA purification kit (Qiagen, Santa Clarita, CA 91355) according to the manufacturer's protocol. First-strand cDNA was synthesized using the NotI/Oligo-dT18 primer. Double-stranded cDNA was synthesized, ligated with EcoRI/BstX I
25 adaptors (Invitrogen, Carlsbad, CA) and digested with NotI. Following size fractionation with Chroma Spin-1000 columns (Clontech, Palo Alto, CA 94303), the cDNA was ligated into the EcoRI/NotI site of pCDNA3.1 (Invitrogen, Carlsbad, CA) and transformed into ElectroMax *E. coli* DH10B cells (BRL Life Technologies) by electroporation.

Using the same procedure, a normal human breast cDNA expression library was prepared from a pool of four normal breast tissue specimens. The cDNA libraries were characterized by determining the number of independent colonies, the percentage of clones that carried insert, the average insert size and by sequence analysis. The breast tumor library contained 1.14×10^7 independent colonies, with more than 90% of clones having a visible insert and the average insert size being 936 base pairs. The normal breast cDNA library contained 6×10^6 independent colonies, with 83% of clones having inserts and the average insert size being 1015 base pairs. Sequencing analysis showed both libraries to contain good complex cDNA clones that were synthesized from mRNA, with minimal rRNA and mitochondrial DNA contamination sequencing.

cDNA library subtraction was performed using the above breast tumor and normal breast cDNA libraries, as described by Hara *et al.* (*Blood*, 84:189-199, 1994) with some modifications. Specifically, a breast tumor-specific subtracted cDNA library was generated as follows. Normal breast cDNA library (70 µg) was digested with EcoRI, NotI, and SfI, followed by a filling-in reaction with DNA polymerase Klenow fragment. After phenol-chloroform extraction and ethanol precipitation, the DNA was dissolved in 100 µl of H₂O, heat-denatured and mixed with 100 µl (100 µg) of Photoprobe biotin (Vector Laboratories, Burlingame, CA), the resulting mixture was irradiated with a 270 W sunlamp on ice for 20 minutes. Additional Photoprobe biotin (50 µl) was added and the biotinylation reaction was repeated. After extraction with butanol five times, the DNA was ethanol-precipitated and dissolved in 23 µl H₂O to form the driver DNA.

To form the tracer DNA, 10 µg breast tumor cDNA library was digested with BamHI and XhoI, phenol chloroform extracted and passed through Chroma spin-400 columns (Clontech). Following ethanol precipitation, the tracer DNA was dissolved in 5 µl H₂O. Tracer DNA was mixed with 15 µl driver DNA and 20 µl of 2 x hybridization buffer (1.5 M NaCl/10 mM EDTA/50 mM HEPES pH 7.5/0.2% sodium dodecyl sulfate), overlaid with mineral oil, and heat-denatured completely. The sample was immediately transferred into a 68 °C water bath and

incubated for 20 hours (long hybridization [LH]). The reaction mixture was then subjected to a streptavidin treatment followed by phenol/chloroform extraction. This process was repeated three more times. Subtracted DNA was precipitated, dissolved in 12 µl H₂O, mixed with 8 µl driver DNA and 20 µl of 2 x hybridization buffer, and
5 subjected to a hybridization at 68 °C for 2 hours (short hybridization [SH]). After removal of biotinylated double-stranded DNA, subtracted cDNA was ligated into BamHI/XhoI site of chloramphenicol resistant pBCSK⁺ (Stratagene, La Jolla, CA 92037) and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate a breast tumor specific subtracted cDNA library.

10 To analyze the subtracted cDNA library, plasmid DNA was prepared from 100 independent clones, randomly picked from the subtracted breast tumor specific library and characterized by DNA sequencing with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A (Foster City, CA). Thirty-eight distinct cDNA clones were found in the subtracted breast tumor-specific cDNA
15 library. The determined 3' cDNA sequences for 14 of these clones are provided in SEQ ID NO: 1-14, with the corresponding 5' cDNA sequences being provided in SEQ ID NO: 15-28, respectively. The determined one strand (5' or 3') cDNA sequences for the remaining clones are provided in SEQ ID NO: 29-52. Comparison of these cDNA sequences with known sequences in the gene bank using the EMBL
20 and GenBank databases (Release 97) revealed no significant homologies to the sequences provided in SEQ ID NO: 3, 10, 17, 24 and 45-52. The sequences provided in SEQ ID NO: 1, 2, 4-9, 11-16, 18-23, 25-41, 43 and 44 were found to show at least some degree of homology to known human genes. The sequence of SEQ ID NO: 42
25 was found to show some homology to a known yeast gene.

25 cDNA clones isolated in the breast subtraction described above were colony PCR amplified and their mRNA expression levels in breast tumor, normal breast and various other normal tissues were determined using microarray technology (Synteni, Fremont, CA). Briefly, the PCR amplification products were dotted onto slides in an array format, with each product occupying a unique location in the array.
30 mRNA was extracted from the tissue sample to be tested, reverse transcribed, and

fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured. This intensity correlates with the hybridization intensity.

Data was analyzed using GEMTOOLS Software. Twenty one distinct
5 cDNA clones were found to be over-expressed in breast tumor and expressed at low levels in all normal tissues tested. The determined partial cDNA sequences for these clones are provided in SEQ ID NO: 53-73. Comparison of the sequences of SEQ ID NO: 53, 54 and 68-71 with those in the gene bank as described above, revealed some homology to previously identified human genes. No significant homologies were
10 found to the sequences of SEQ ID NO: 55-67, 72 (referred to as JJ 9434) and 73 (referred to as B535S). In further studies, full length cDNA sequences were obtained for the clones 1016F8 (SEQ ID NO: 56; also referred to as B511S) and 1016D12 (SEQ ID NO: 61; also referred to as B532S), and an extended cDNA sequence was obtained for 1012H8 (SEQ ID NO: 64; also referred to as B533S). These cDNA
15 sequences are provided in SEQ ID NO: 95-97, respectively, with the corresponding predicted amino acid sequences for B511S and B532S being provided in SEQ ID NO: 98 and 99, respectively.

Analysis of the expression of B511S in breast tumor tissues and in a variety of normal tissues (skin, PBMC, intestine, breast, stomach, liver, kidney, fetal
20 tissue, adrenal gland, salivary gland, spinal cord, large intestine, small intestine, bone marrow, brain, heart, colon and pancreas) by microarray, northern analysis and real time PCR, demonstrated that B511S is over-expressed in breast tumors, and normal breast, skin and salivary gland, with expression being low or undetectable in all other tissues tested.

25 Analysis of the expression of B532S in breast tumor tissue and in a variety of normal tissues (breast, PBMC, esophagus, HMEC, spinal cord, bone, thymus, brain, bladder, colon, liver, lung, skin, small intestine, stomach, skeletal muscle, pancreas, aorta, heart, spleen, kidney, salivary gland, bone marrow and adrenal gland) by microarray, Northern analysis and real time PCR, demonstrated that

B532S is over-expressed in 20-30% of breast tumors with expression being low or undetectable in all other tissues tested.

In a further experiment, cDNA fragments were obtained from two subtraction libraries derived by conventional subtraction, as described above and 5 analyzed by DNA microarray. In one instance the tester was derived from primary breast tumors, referred to as Breast Subtraction 2, or BS2. In the second instance, a metastatic breast tumor was employed as the tester, referred to as Breast Subtraction 3, or BS3. Drivers consisted of normal breast.

cDNA fragments from these two libraries were submitted as templates 10 for DNA microarray analysis, as described above. DNA chips were analyzed by hybridizing with fluorescent probes derived from mRNA from both tumor and normal tissues. Analysis of the data was accomplished by creating three groups from the sets of probes, referred to as breast tumor/mets, normal non-breast tissues, and metastatic breast tumors. Two comparisons were performed using the modified Gemtools 15 analysis. The first comparison was to identify templates with elevated expression in breast tumors. The second was to identify templates not recovered in the first comparison that yielded elevated expression in metastatic breast tumors. An arbitrary level of increased expression (mean of tumor expression versus the mean of normal tissue expression) was set at approximately 2.2.

20 In the first round of comparison to identify over-expression in breast tumors, two novel gene sequences were identified, hereinafter referred to as B534S and B538S (SEQ ID NO: 89 and 90, respectively), together with six sequences that showed some degree of homology to previously identified genes (SEQ ID NO: 74-79). The sequences of SEQ ID NO: 75 and 76 were subsequently determined to be 25 portions of B535S (SEQ ID NO: 73). In a second comparison to identify elevated expression in metastatic breast tumors, five novel sequences were identified, hereinafter referred to as B535S, B542S, B543S, P501S and B541S (SEQ ID NO: 73 and 91-94, respectively), as well as nine gene sequences that showed some homology to known genes (SEQ ID NO: 80-88). Clone B534S and B538S (SEQ ID NO: 89 and

90) were shown to be over-expressed in both breast tumors and metastatic breast tumors.

In a subsequent series of studies, 457 clones from Breast Subtraction 2 were analyzed by microarray on Breast Chip 3. As described above, a first comparison to identify over-expression in breast tumors over normal non-breast tissues was performed. This analysis yielded six cDNA clones that demonstrated elevated expression in breast tumor over normal non-breast tissues. Two of these clones, referred to as 1017C2 (SEQ ID NO: 102) and B546S (SEQ ID NO: 107) do not share significant homology to any known genes. Clone B511S also showed over-expression in breast tumor, which was previously described as 1016F8, with the determined cDNA sequence provided in SEQ ID NO: 95 and the predicted amino acid sequence provided in SEQ ID NO: 98. The remaining four clones over-expressed in breast tumor were found to share some degree of homology to Tumor Expression Enhanced Gene (SEQ ID NO: 103 and 104) Stromelysin-3 (SEQ ID NO: 105) or Collagen (SEQ ID NO: 106).

In the second comparison to determine genes with elevated expression in metastatic breast tumors over non-breast normal tissues, a profile similar to the first comparison was derived. The two putatively novel clones, 1017C2 and B546S, SEQ ID NO: 102 and 107, respectively, were overexpressed in metastatic breast tumors. In addition, Tumor Expression Enhanced Gene and B511S also showed elevated expression in metastatic breast tumors.

As described in U.S. Patent Application No. 08/806,099, filed February 25, 1997, the antigen P501S was isolated by subtracting a prostate tumor cDNA library with a normal pancreas cDNA library and with three genes found to be abundant in a previously subtracted prostate tumor specific cDNA library: human glandular kallikrein, prostate specific antigen (PSA), and mitochondria cytochrome C oxidase subunit II. The determined full-length cDNA sequence for P501S is provided in SEQ ID NO: 100, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 101. Expression of P501S in breast tumor was examined by microarray analysis. Over-expression was found in prostate tumor, breast tumor

and metastatic breast tumor, with negligible to low expression being seen in normal tissues. This data suggests that P501S may be over-expressed in various breast tumors as well as in prostate tumors.

5

Example 2

GENERATION OF HUMAN CD8+ CYTOTOXIC T-CELLS THAT RECOGNIZE ANTIGEN PRESENTING CELLS EXPRESSING BREAST TUMOR ANTIGENS

- 10 This Example illustrates the generation of T cells that recognize target cells expressing the antigen B511S, also known as 1016-F8 (SEQ ID NO: 56). Human CD8+ T cells were primed *in-vitro* to the B511S gene product using dendritic cells infected with a recombinant vaccinia virus engineered to express B511S as follows (also see Yee et al., Journal of Immunology (1996) 157 (9):4079-86).
- 15 Dendritic cells (DC) were generated from peripheral blood derived monocytes by differentiation for 5 days in the presence of 50 µg/ml GMCSF and 30 µg/ml IL-4. DC were harvested, plated in wells of a 24-well plate at a density of 2×10^5 cells/well and infected for 12 hours with B511S expressing vaccinia at a multiplicity of infection of 5. DC were then matured overnight by the addition of 3 µg/ml CD40-
- 20 Ligand and UV irradiated at 100µW for 10 minutes. CD8+ T cells were isolated using magnetic beads, and priming cultures were initiated in individual wells (typically in 24 wells of a 24-well plate) using 7×10^5 CD8+ T cells and 1×10^6 irradiated CD8-depleted PBMC. IL-7 at 10 ng/ml was added to cultures at day 1. Cultures were re-stimulated every 7-10 days using autologous primary fibroblasts
- 25 retrovirally transduced with B511S and the costimulatory molecule B7.1. Cultures were supplemented at day 1 with 15 I.U. of IL-2. Following 4 such stimulation cycles, CD8+ cultures were tested for their ability to specifically recognize autologous fibroblasts transduced with B511S using an interferon-γ Elispot assay (see Lalvani et al J. Experimental Medicine (1997) 186:859-965). Briefly, T cells from
- 30 individual microcultures were added to 96-well Elispot plates that contained autologous fibroblasts transduced to express either B511S or as a negative control

- antigen EGFP, and incubated overnight at 37° C; wells also contained IL-12 at 10 ng/ml. Cultures were identified that specifically produced interferon- γ only in response to B511S transduced fibroblasts; such lines were further expanded and also cloned by limiting dilution on autologous B-LCL retrovirally transduced with B511S.
- 5 Lines and clones were identified that could specifically recognize autologous B-LCL transduced with B511S but not autologous B-LCL transduced with the control antigens EGFP or HLA-A3. An example demonstrating the ability of human CTL cell lines derived from such experiments to specifically recognize and lyse B511S expressing targets is presented in Figure 1.

10

Example 3

SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on an Perkin Elmer/Applied Biosystems Division 430A peptide synthesizer using Fmoc chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for the purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention.

CLAIMS

1. An isolated polypeptide comprising an immunogenic portion of a breast protein or a variant thereof, wherein said protein comprises an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of: (a) nucleotide sequences recited in SEQ ID NOS: 3, 10, 17, 24, 45-52, 55-67, 72, 73, 89-97, 102 and 107; (b) complements of said nucleotide sequences; and (c) sequences that hybridize to a sequence of (a) or (b) under moderately stringent conditions.

2. The isolated polypeptide of claim 1, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 98, 99 and 101.

3. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of any one of claims 1 and 2.

4. An isolated polynucleotide comprising a sequence provided in SEQ ID NOS: 3, 10, 17, 24, 45-52, 55-67, 72, 73, 89-97, 102 and 107.

5. An expression vector comprising a polynucleotide according to any one of claims 3 and 4.

6. A host cell transformed with the expression vector of claim 5.

7. The host cell of claim 6 wherein the host cell is selected from the group consisting of *E. coli*, yeast and mammalian cell lines.

8. A pharmaceutical composition comprising the polypeptide of claim 1 and a physiologically acceptable carrier.

9. A vaccine comprising the polypeptide of claim 1 and a non-specific immune response enhancer.

10. The vaccine of claim 9 wherein the non-specific immune response enhancer is an adjuvant.

11. A vaccine comprising an isolated polynucleotide of any one of claims 3 and 4, and a non-specific immune response enhancer.

12. The vaccine of claim 11 wherein the non-specific immune response enhancer is an adjuvant.

13. A pharmaceutical composition for the treatment of breast cancer comprising a polypeptide and a physiologically acceptable carrier, the polypeptide comprising an immunogenic portion of a breast protein, wherein said protein comprises an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of: (a) nucleotide sequences recited in SEQ ID NOS: 1, 2, 4-9, 11-16, 18-23, 25-44, 53, 54, 68-71, 74-88 and 103-106; (b) complements of said nucleotide sequences; and (c) sequences that hybridize to a sequence of (a) or (b) under moderately stringent conditions.

14. A vaccine for the treatment of breast cancer comprising a polypeptide and a non-specific immune response enhancer, said polypeptide comprising an immunogenic portion of a breast protein, wherein said protein comprises an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of: (a) nucleotide sequences recited in SEQ ID NOS: 1, 2, 4-9, 11-16, 18-23, 25-44, 53, 54, 68-71, 74-88 and 103-106; (b) complements of said nucleotide sequences; and (c) sequences that hybridize to a sequence of (a) or (b) under moderately stringent conditions.

15. The vaccine of claim 14 wherein the non-specific immune response enhancer is an adjuvant.

16. A vaccine for the treatment of breast cancer comprising a polynucleotide and a non-specific immune response enhancer, the polynucleotide

comprising a sequence selected from the group consisting of: (a) nucleotide sequences recited in SEQ ID NOS: 1, 2, 4-9, 11-16, 18-23, 25-44, 53, 54, 68-71, 74-88 and 103-106; (b) complements of said nucleotide sequences; and (c) sequences that hybridize to a sequence of (a) or (b) under moderately stringent conditions.

17. The vaccine of claim 16, wherein the non-specific immune response enhancer is an adjuvant.

18. A pharmaceutical composition according to any one of claims 8 and 13, for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

19. A vaccine according to any one of claims 9, 11, 14 or 16, for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

20. A fusion protein comprising at least one polypeptide according to claim 1.

21. A pharmaceutical composition comprising a fusion protein according to claim 20 and a physiologically acceptable carrier.

22. A vaccine comprising a fusion protein according to claim 20 and a non-specific immune response enhancer.

23. The vaccine of claim 22 wherein the non-specific immune response enhancer is an adjuvant.

24. A pharmaceutical composition according to claim 21, for use in manufacture of a medicament for inhibiting the development of breast cancer in a patient.

25. A vaccine according to claim 22, for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

26. A method for detecting breast cancer in a patient, comprising:

(a) contacting a biological sample from a patient with a binding agent which is capable of binding to a polypeptide, the polypeptide comprising an immunogenic portion of a breast protein, wherein said protein comprises an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of nucleotide sequences recited in SEQ ID NOS: 1-97, 100 and 102-107, complements of said nucleotide sequences and sequences that hybridize to a sequence provided in SEQ ID NO: 1-97, 100 and 102-107 under moderately stringent conditions; and

(b) detecting in the sample a protein or polypeptide that binds to the binding agent, thereby detecting breast cancer in the patient.

27. The method of claim 26 wherein the binding agent is a monoclonal antibody.

28. The method of claim 27 wherein the binding agent is a polyclonal antibody.

29. A method for monitoring the progression of breast cancer in a patient, comprising:

(a) contacting a biological sample from a patient with a binding agent that is capable of binding to a polypeptide, said polypeptide comprising an immunogenic portion of a breast protein, wherein said protein comprises an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of nucleotide sequences recited in SEQ ID NOS: 1-97, 100 and 102-107, complements of said nucleotide sequences and sequences that hybridize to a sequence provided in SEQ ID NO: 1-97, 100 and 102-107 under moderately stringent conditions;

(b) determining in the sample an amount of a protein or polypeptide that binds to the binding agent;

- (c) repeating steps (a) and (b); and
- (d) comparing the amount of polypeptide detected in steps (b) and (c)

to monitor the progression of breast cancer in the patient.

30. A monoclonal antibody that binds to a polypeptide comprising an immunogenic portion of a breast protein or a variant of said protein that differs only in conservative substitutions and/or modifications, wherein said protein comprises an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of: (a) nucleotide sequences recited in SEQ ID NOS: 3, 10, 17, 24, 45-52, 55-67, 72, 73, 89-97, 102 and 107; (b) complements of said nucleotide sequences; and (c) sequences that hybridize to a sequence of (a) or (b) under moderately stringent conditions.

31. A monoclonal antibody according to claim 30, for use in the manufacture of a medicament for inhibiting the development of breast cancer in a patient.

32. The monoclonal antibody of claim 31 wherein the monoclonal antibody is conjugated to a therapeutic agent.

33. A method for detecting breast cancer in a patient comprising:

- (a) contacting a biological sample from a patient with at least two oligonucleotide primers in a polymerase chain reaction, wherein at least one of the oligonucleotides is specific for a polynucleotide encoding a polypeptide comprising an immunogenic portion of a breast protein, said protein comprising an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of nucleotide sequences recited in SEQ ID NO: 1-97, 100 and 102-107, complements of said nucleotide sequences and sequences that hybridize to a sequence of SEQ ID NO: 1-97, 100 and 102-107 under moderately stringent conditions; and
- (b) detecting in the sample a polynucleotide sequence that amplifies in the presence of the oligonucleotide primers, thereby detecting breast cancer.

34. The method of claim 33, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotides of a polynucleotide comprising a sequence selected from SEQ ID NOS: 1-97, 100 and 102-107.

35. A diagnostic kit comprising:

- (a) one or more monoclonal antibodies of claim 30; and
- (b) a detection reagent.

36. A diagnostic kit comprising:

(a) one or more monoclonal antibodies that bind to a polypeptide encoded by a polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NOS: 1, 2, 4-9, 11-16, 18-23, 25-44, 53, 54, 68-71, 74-88 and 103-106, complements of said sequences, and sequences that hybridize to a sequence of SEQ ID NO: 1, 2, 4-9, 11-16, 18-23, 25-44, 53, 54, 68-71, 74-88 or 103-106 under moderately stringent conditions; and

- (b) a detection reagent.

37. The kit of claims 35 or 36 wherein the monoclonal antibodies are immobilized on a solid support.

38. The kit of claim 37 wherein the solid support comprises nitrocellulose, latex or a plastic material.

39. The kit of claims 35 or 36 wherein the detection reagent comprises a reporter group conjugated to a binding agent.

40. The kit of claim 39 wherein the binding agent is selected from the group consisting of anti-immunoglobulins, Protein G, Protein A and lectins.

41. The kit of claim 39 wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.

42. A diagnostic kit comprising at least two oligonucleotide primers, at least one of the oligonucleotide primers being specific for a polynucleotide encoding a polypeptide comprising an immunogenic portion of a breast protein, said protein comprising an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of nucleotide sequences recited in SEQ ID NOS: 1-97, 100 and 102-107 complements of said nucleotide sequences and sequences that hybridize to a sequence of SEQ ID NO: 1-97, 100 and 102-107 under moderately stringent conditions.

43. A diagnostic kit of claim 42 wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotides of a polynucleotide comprising a sequence selected from SEQ ID NOS: 1-97, 100 and 102-107.

44. A method for detecting breast cancer in a patient, comprising:

- (a) obtaining a biological sample from the patient;
- (b) contacting the biological sample with an oligonucleotide probe specific for a polynucleotide encoding a polypeptide comprising an immunogenic portion of a breast protein, said protein comprising an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of nucleotide sequences recited in SEQ ID NOS: 1-97, 100 and 102-107, complements of said nucleotide sequences and sequences that hybridize to a sequence of SEQ ID NO: 1-97, 100 and 102-107 under moderately stringent conditions; and
- (c) detecting in the sample a polynucleotide sequence that hybridizes to the oligonucleotide probe, thereby detecting breast cancer in the patient.

45. The method of claim 44 wherein the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a polynucleotide comprising a sequence selected from the group consisting of SEQ ID NOS: 1-97, 100 and 102-107.

46. A diagnostic kit comprising an oligonucleotide probe specific for a polynucleotide encoding a polypeptide comprising an immunogenic portion of a breast protein, said protein comprising an amino acid sequence encoded by a polynucleotide comprising a sequence selected from the group consisting of nucleotide sequences recited in SEQ ID NOS: 1-97, 100 and 102-107, complements of said nucleotide sequences, and sequences that hybridize to a sequence of SEQ ID NO: 1-97, 100 and 102-107 under moderately stringent conditions.

47. The diagnostic kit of claim 46, wherein the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a polynucleotide comprising a sequence selected from the group consisting of SEQ ID NOS: 1-97, 100 and 102-107.

48. A method for treating breast cancer in a patient, comprising the steps of:

- (a) obtaining peripheral blood cells from the patient;
- (b) incubating the cells in the presence of at least one polypeptide of any one of claims 1 and 2, such that T cells proliferate; and administering the proliferated T cells to the patient.

49. A method for treating breast cancer in a patient, comprising the steps of:

- (a) obtaining peripheral blood cells from the patient;
- (b) incubating the cells in the presence of at least one polynucleotide of any one of claims 3 and 4, such that T cells proliferate; and
- (c) administering the proliferated T cells to the patient.

50. The method of any one of claims 48 and 49 wherein the step of incubating the cells is repeated one or more times.

51. The method of any one of claims 48 and 49 wherein step (a) further comprises separating the T cells from the peripheral blood cells and the cells incubated in step (b) are the T cells.

52. The method of any one of claims 48 and 49 wherein step (a) further comprises separating CD4+ cells or CD8+ cells from the peripheral blood cells and the cells proliferated in step (b) are CD4+ or CD8+ T cells.

53. The method of any one of claims 48 and 49 wherein step (b) further comprises cloning at least one T cell that proliferated in the presence of the polypeptide.

54. A composition for the treatment of breast cancer in a patient, comprising T cells proliferated in the presence of a polypeptide of any one of claims 1 and 2, in combination with a pharmaceutically acceptable carrier.

55. A composition for the treatment of breast cancer in a patient comprising T cells proliferated in the presence of a polynucleotide of any one of claims 3 and 4, in combination with a pharmaceutically acceptable carrier.

56. A method for treating breast cancer in a patient, comprising the steps of:

- (a) incubating antigen presenting cells in the presence of at least one polypeptide of any one of claims 1 and 2; and
- (b) administering to the patient the incubated antigen presenting cells.

57. A method for treating breast cancer in a patient, comprising the steps of:

- (a) incubating antigen presenting cells in the presence of at least one polynucleotide of any one of claims 3 and 4; and
- (b) administering to the patient the incubated antigen presenting

cells.

58. The method of claims 56 or 57 wherein the antigen presenting cells are selected from the group consisting of dendritic cells, macrophage cells, monocyte cells, fibroblast cells, B-cells or combinations thereof.

59. A composition for the treatment of breast cancer in a patient, comprising antigen presenting cells incubated in the presence of a polypeptide of any one of claims 1 and 2, in combination with a pharmaceutically acceptable carrier.

60. A composition for the treatment of breast cancer in a patient, comprising antigen presenting cells incubated in the presence of a polynucleotide of any one of claims 3 and 4, in combination with a pharmaceutically acceptable carrier.

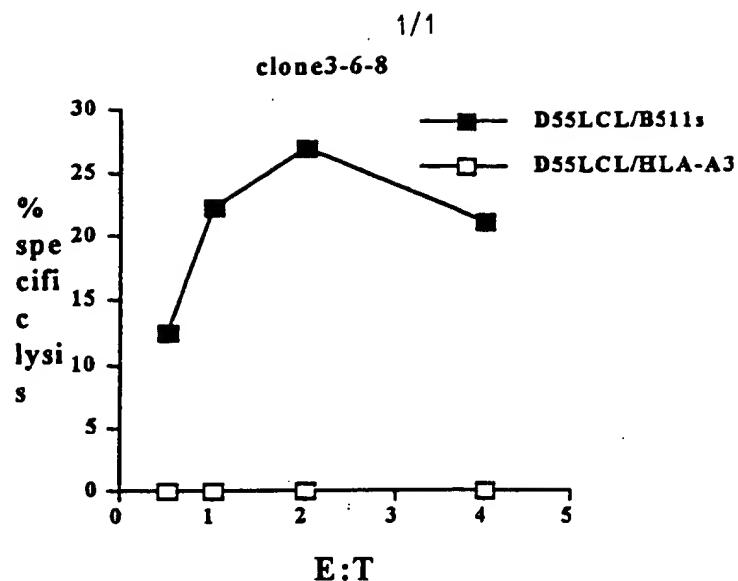
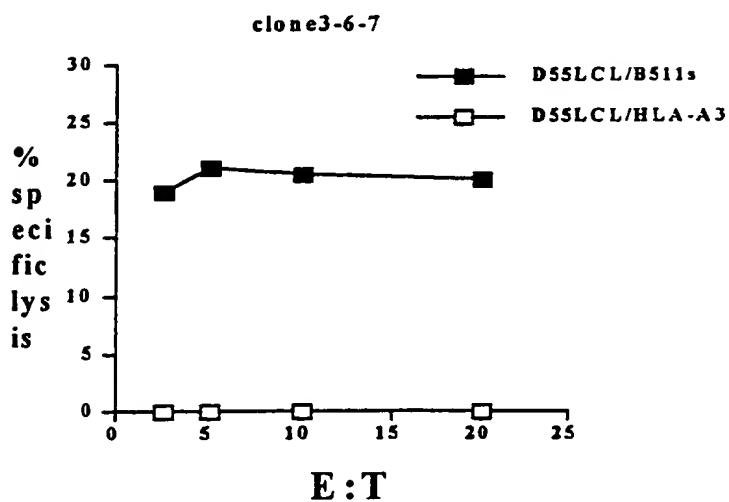
**FIGURE 1A****FIGURE 1B**

Figure 1: Specific lytic activity of B511s-specific CTL clones 3-6-8 and 3-6-7 measured on autologous LCL transduced with B511s (filled squares) or HLA-A3 (open squares). Each data point is the average of triplicate measurements.

FIGS. 1A AND 1B

SEQUENCE LISTING

<110> Corixa Corporation
Reed, Steven G.
Xu, Jiangchun
Dillon, Davin C.

<120> COMPOUNDS FOR IMMUNOTHERAPY AND
DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE

<130> 210121.44602PC

<140> PCT
<141> 2000-04-10

<160> 107

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 402
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(402)
<223> n = A,T,C or G

<400> 1

tttttttttt	tttttaggag	aactgaatca	aacagatttt	attcaacttt	tttagatgagg	60
aaaacaaatn	atacgaaatn	ngtcataaga	aatgccttct	tataccacta	tctcaaacca	120
ctttcaatat	tttacaaaaat	gctcacgcag	caaatatgaa	aagctncaac	acttcccttt	180
gttaacttgc	tgcaatnaat	gcaactttaa	canacataca	aatttcttct	gtatcttaaa	240
agttnaatta	ctaattttaa	tgatnttnt	caagatnttt	attcatatac	ttttaatgac	300
tcnttgcncna	tacatacnnta	ttttcttac	tttttttta	cnatnggccca	acagcttca	360
ngcagnccnc	aaaaatctta	ccggtaatt	acacggggtt	gt		402

<210> 2

<211> 424
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(424)
<223> n = A,T,C or G

<400> 2

tttttttttt	tttttaaag	gtacacattt	cttttcatt	ctgtttnatg	cagcaaataa	60
ttcggtggca	tcttctctgt	gatgggcagc	ttgctaaaat	tanactcagg	cccccttagct	120
ncatttccaa	ctnagcccac	gctttcaacc	nngccnaaca	aagaaaatca	gttngggtta	180
aattcttgc	tgganacaaa	gaactacatt	cctttgtaaa	tnatgctttg	tttgctctgt	240

gcaaacncag attgaaggga anaagganac ttntgggac ggaaacaact ngnagaagca	300
gganccgccc aggnccattt cctcaccatg cttaatcttg cnctcacttg cnnggcacca	360
ttaaacttgg tgcaaaaaggc gcaattgtg nanggaaccc cacaccttc ttaaaaagca	420
gggc	424
<210> 3	
<211> 421	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(421)	
<223> n = A,T,C or G	
<400> 3	
tttttttttt ttttcccaa tttaaaaaag ccttttcat acttcaatta caccanactt	60
aatnatttca tgtagtaatc ngacattatt atttnaaaat ttgcataattt aaaatttgna	120
tcanttactt ccagactgtt tgcanaatga agggaggatc actcaagngc tgatctcnca	180
ctntctgcag tctnctgtcc ttgtccccgn ctaatggatc gacactanat ggacagntcn	240
cagatcttcc gttcttntcc cttcccaat ttncnccanc tcccccctt ncccgatcn	300
tttggggaca tgnataattt gcnatcccta aaccctgccc gccangggtc ccnanctcag	360
gggtgggtaa tgttcgnncng gcttnttgac ncctgcgcc ctttnantcc naaccccaag	420
c	421
<210> 4	
<211> 423	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(423)	
<223> n = A,T,C or G	
<400> 4	
tttttttatt ttttttcta tttntnntat ttnntgnggt tcctgtgtgt aattagnang	60
tgtgtatgcg tangtacnta tggtnatggc agttaanaga gactatngat naattttAAC	120
aaaatctcaa natngtattt ggttnatggc agttaanaga gactatngat naattttAAC	180
atggacacng tggaaatgttag cgcctnatca nttaaaaact tcattttgaa ggcctttnc	240
cctccnaata aaaatnccng gcccactgg gttaagcaac attgcatttc taaagaaacc	300
acatgcanac nagttaaacc tggtnactgg tcangcaaac cnanntggaa nanaagggn	360
ttcncccan ggacantcng aatttttta acaaattacn atncccccc nffffggagcc	420
tgt	423
<210> 5	
<211> 355	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(355)	
<223> n = A,T,C or G	

```

<400> 5
acgaccacct natttcgtat ctttcaactc tttcgaccg gacctcttat tcggaagcgt      60
tccaggaaga caggtctcaa ctttagggatc agatcacgtt atcaacgctc tgggatcgct    120
gcaacctggc acttcaagga agtgcaccga tnacgtctag accggccaac acagatctag    180
aggtggccaa ctgatcaactg taggagctga ctggcaanan tcaaccgggc cccaaaccnag   240
agtgaccaan acnaccatn aggatcaccc acaggcactc ctcgtcctag ggccaaccnna   300
ccaaacggct ggccaatggg gggggttaat atttggttna aaaattgatt ttaaa            355

<210> 6
<211> 423
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(423)
<223> n = A,T,C or G

<400> 6
ttttttttt ttttggaca ggaagtaaaa tttattggtn antattaana ggggggcagc      60
acatttggaaag ccctcatgan tgcagggccc gccacttgc cagagggcca cnattgggaa    120
tgtacttaac cccacagccn tctggatna gccgctttc agccaccatn tcttcaaatt     180
catcagcatt aaacttggta aanccccact tctttaagat ntgnatctc tggccggccag    240
naaaacttggaa cttggccctg cgccaggccct caatcacatg ctcccttgc tgcagcttgg  300
tgcgnaagga cntaatnact tggccnatgt gaaccctggc cacantggcc tgggctttc   360
caaaggcacc tgcgaaggcct ntttggancc tgnccgcccc ngcacagggaa caacatctt   420
ttt                                         423

<210> 7
<211> 410
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(410)
<223> n = A,T,C or G

<400> 7
ttcgcaactgg ctaaaacaaa ccgccttgca aagtngaaa aatttatcaa tggaccaaat      60
aatgctcata tccnacaagt tggtgaccgt ntnttnata aaaaaatgtt tnatgctcct    120
nanttgggtt acaataatgt tccaaattng gacnttcggc atctaccctg gttcacctgg   180
gtaaatatca ggcagctttt gatggggcta ggaaagctaa cagtactcga acatggggaaa  240
gaggtctgct tcgccngtgt anatggggaaa naattccgtc ttgctcngat ttgtggactt  300
catattgtt tacatgcaga tgaatnngaa gaacttgtca actactatca ggatcgtggc   360
tttttnaaaa agctnatcac catgttggaa gcggcactng gacttgagcg             410

<210> 8
<211> 274
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(274)

```

<223> n = A,T,C or G

<400> 8

ttttttttt	tttttaggtc	atacatat	tttattataa	canatatn	tatatacata	60
taatataat	gtatatatcc	acgtgtgt	gtgtgtatca	aaaacaacan	aanttagtg	120
atctatata	ntngctcaca	tatgcatt	ggg agataccagt	aaaaaataag	tnaatctcca	180
taatatgtt	taaaactcan	anaaaatcnga	gagactnaaa	gaaaacgtt	atcannatga	240
tttgtngataa	tcttgaan	aa tnacnaaaac	at			274

<210> 9

<211> 322

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(322)

<223> n = A,T,C or G

<400> 9

ttttttttt	tttgcac	gcnaaactt	ctagcactat	attaaactca	60
ataagagtga	taagtgtgaa	aatcctg	ttctctttaa	tcttaatgna	120
gttttcacc	attaantgt	ataatggct	tatgtat	tatnnatgg	180
ttaaaaaaagt	tttcctct	ccctngtt	ctaanagtt	tatcaaaaa	240
attnngttca	gtacttttnc	ctgcac	ctat agatatgatn	tggtataat	300
cctnnanata	tgatggatna	ca			322

<210> 10

<211> 425

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(425)

<223> n = A,T,C or G

<400> 10

ttttttttt	tttttattct	gcagccatta	aatgctgaac	actagatn	tatttg	60
ggtcacaaaa	taagtacaga	attnacaca	cgccctgccc	ataaaaagca	cagctcccag	120
ttctatattt	acaatatctc	tggat	cctcccttc	taattt	gact aatattctg	180
cttctcaggc	agcagcgc	cctggcaacc	ataagaacca	acntgng	gac taggtcggt	240
ggccaaggat	caggaaacag	aanaatggaa	gnagcccccn	tgacnctatt	aanctntnaa	300
actatctnaa	ctgctagttt	tcaggctt	aaatcatgtaa	natacgt	cttnttgctg	360
caaccggaag	cattcctagat	ggtacact	ctccaggtgc	caggaaaaga	tcccaaatng	420
caggn						425

<210> 11

<211> 424

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(424)

```

<223> n = A,T,C or G

<400> 11
ttttnttant ttttttancc nctnntccnn tntgtgnag ggggtaccaa atttctttat
ttaaagaat ggtacaaatc aaaaaactta atttaattt tnggtacaac ttatagaaaa
120ggtaaggaa accccaacat gcatgcactg ccttggtaac cagggnattc cccncggct
180
ntggggaaat tagcccaang ctnagcttc attatcactn tcccccaggg tntgctttc
aaaaaaaaattt nccgcnagn cnaatccggg cnctccatc tggcgcaant tggtcacttg
gtccccnat tcttaangg cttncacctn ctcattcggg tnatgtgtct caattaaatc
ccacngatgg gggtcattt tntcnntag ccagtttgtg nagttccgtt attganaaaa
ccan

<210> 12
<211> 426
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(426)
<223> n = A,T,C or G

<400> 12
tttttttttt ttttncttaa aagctttat ctcctgctta cattaccat ctgttcttgc
atgttgctg cttttccac tagagccctt aacaacttaa tcatggttat tttaagggt
ctaataattc cnaaactggt atcataaata agtctcggtt tnatgtgtt tttctctcta
tcacactgtg ttngttgctt ttnacatgc tttgtatatt ttggctgaaa gctgaaaaat
nacatacctg gtnttacaac ctgaggtaan cagcctnta gtgtgaggtt ttatatntta
ctggctaaga gctnggcnt gttnantant tgggtgact ntatatgcca naggcttnta
tttccnctng tgcccttgc ttagtacccc attnttttag gggttcccta naaactctat
ctnaat

<210> 13
<211> 419
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(419)
<223> n = A,T,C or G

<400> 13
tttttttttt ttttnagat agactctcac tctttcgccc aggctggagt gcagtgccgc
aatcaaggct cactgcaacc tctgccttat aaagcattt ctaaagggtac aagctaaatt
ttaaaaatctctncacaa ctaatgtata acaaaaatta gttctacccataaaacncnt
ggctcagccc tcgnaacaca tttccctgtt ctcaactgtat gaacactcca naaacagaac
anatntaagc tttccaggc ccagaaaagc tcgctggggg atttgctntg tgggtgacac
acttgccacc ctgtggcagc acagctccac acntgctttg ggccgcattt gcaagttctc
tgtaancccc ctgnaagacc cggatcagct gggtnagaaat tgcangcnc ttttggca

<210> 14
<211> 400
<212> DNA

```

```

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(400)
<223> n = A,T,C or G

<400> 14
aanccattgc caagggtatc cggaggattg tggctgtcac aggtnccgag gcccanaagg      60
ccctcaggaa agcaaagagc ttgaaaatg tctctctgtc atggaagccn aagtgaaggc      120
tcanactgct ccaacaagga tntgcanagg gagatcgcta accttggaga ggccctggcc      180
actgcagtcn tcccccantg gcagaaggat gaattgcggg agactctcan atcccttang      240
gaaggtcgtg gatnactttg accgagcctc nnaagccaat ntccagaaca agtgttggag      300
aagacaaaagc anttcatcga cgccaaacccc naccggcctc tnttctcctg ganattgana      360
gcggcgcccc cgccccagggc cttaataanc cntgaagctn      400

<210> 15
<211> 395
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(395)
<223> n = A,T,C or G

<400> 15
tgctttgctg cgtccaggaa gattagatng aanaatacat attgatttgc caaatgaaca      60
agcgagatta gacntactga anatccatgc aggtcccatt acaaaggcatg gtgaaataga      120
tgatgaagca attgtgaagc tatcgatgg ctttnatgga gcagatctga gaaatgtttg      180
tactgaagca ggtatgttcg caattcgtgc tgatcatgat tttgttagtac aggaagactt      240
catgaaagcn gtcagaanag tggctnattc tnnaagctgg agtctaaatt ggacnacnac      300
ctntgtatTT actgttggan ttttgatgtc gcatgacaga ttttgcttan tgtaaaaatn      360
aagtcaaga aaattatgtt agttttggcc attat      395

<210> 16
<211> 404
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(404)
<223> n = A,T,C or G

<400> 16
ccaccactaa aatccctggct gagccctacn agtacctgtg cccctcccc aggacgagat      60
nagggcacac ccttaagtn aggtgacagg tcacctttaa gtgaggacag tcagctnaat      120
ttcacctctt gggcttgagt acctgggtct cgtgcccgtga ggcgaacnctn agccctgcag      180
ctnccatgtc cgtgctgccat atngtcttgc tcttctccac gcnctnaac ttggcttca      240
gtaggagctg caggcnagaa ngaagcggtt aacagcgcca ctccatagcc gcagccnggc      300
tgccctgtct tctcaaggag ggggtgtgggg ttccctccacc atcgccgccc ttgcaaacac      360
ntctcanggc ttccctnccg gctnancgca ngacttaagc atgg      404

<210> 17

```

```

<211> 360
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(360)
<223> n = A,T,C or G

<400> 17
ggccagaagc ttccacaaa ccagtgaagg tggcagcaaa gaaaggctct tagacnagga      60
gctggcagca gctgctatct ngatngacng cagaaaccaa ccactaattc agcaaacaca      120
acctcataacc tnaccgcttc cctttnaatg gccttcggtg tgtgcgcaca tggcacgtg      180
cggggagaac catacttatt cccctnttcc cggcctacca cctctnctcc cccttctctt      240
ctctncaatt actntctccn ctgcttntt ctnancacta ctgctngtnt cnanagccng      300
cccgcaatta cctggcaaaa ctcgcgaccc ttccggcagc gctaaanaat gcacatttac      360

<210> 18
<211> 316
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(316)
<223> n = A,T,C or G

<400> 18
atacatatac acatatatga ttttagatag agccatatac ctngaagtag tanatttgg      60
tgtgtgtata tgtagtggtc tactcattt aaataaaactt gtgatagaga tgtaattntg      120
agccagtttt tcatttgctt aaatnactca ccaagtaact aattaaggtn tctttactct      180
taatgttnag tagtgagatt ctgttgaaagg tgatattaaa aaccattcta tattaattaa      240
cattcatgtt gttttttaaa agcttatttgg aaatcnaatt atgattattt ttcataccag      300
tcgatnttat gtangt      316

<210> 19
<211> 350
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(350)
<223> n = A,T,C or G

<400> 19
aaggatgca nataatgctg tgtagtgagct tgatggaaaa gaactctgta gtgaaagggt      60
tactattgaa catgctnggg ctcggcgtacg aggtggaaaga ggttagaggac gataactctga      120
ccgttttagt agtcgcagac ctcgaaatga tagacgaaat gctccacctg taagaacaga      180
anatcgctt atagttgaga atttatccctc aagagtcaac tggcagggtt gttganatac      240
agttttgagt tnttttgatg tggctttta aaaaagttat gggttactna tgttatattg      300
ttttattaaa agtagtttta aattaatgga tntgatggaa ttgttggttt      350

<210> 20
<211> 367

```

```

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(367)
<223> n = A,T,C or G

<400> 20
gntnnncnca agatcctnct ntcccccnng gcngcccnnc cnccngtnat naccgtttn
ntaanatcnn gccgcncccgg aagtctcnct nntgccgaga tgncccttat ncncnnatgn
ncaattntga cctnnggcga anaatggcng nntgttatca gtntccnctc tgnggnctct
tagnatctga ccactangac ccnctatcct ctcaaaccct gtanncngcc ctaatttg
ccaatttagtgcatgntanag cntcctggcc cagatggcnt ccatatcctg gtncggcttc
cgcccctacc angncatccn catctactag agcttatccg ctncntgngg cgacccggnt
ccccnctc 60
120
180
240
300
360
367

<210> 21
<211> 366
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(366)
<223> n = A,T,C or G

<400> 21
cccaacacaa tggtctaagt anaactgtat tgctctgtat tatagtttcca cattggcaac
ctacaatggg aaaatccata cataagtca gttacttcctn atgagctttc tccttctgaa
tcctttatct tctgaagaaa gtacacacct tggtnatgtat atctttgaat tgcccttctt
tccaggcatac agttggatga ttcatcatgg taattatggc attatcatat tcttcatact
tgtcatacga aaacaccagt tctgcccnnna gatgagcttg ttctgcagct cttagcacct
tggaatattt cactctagac cagaaacacgc tcccggtgct ccctcatttt ctgaggctta
aatttn 60
120
180
240
300
360
366

<210> 22
<211> 315
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(315)
<223> n = A,T,C or G

<400> 22
acttaatgca atctctggag gataatttgg atcaagaaat aaagaanaaa tgaatttagga
gaagaaatna ctgggtnata tttcaatatt ttagaacttt aanaatgttg actatgattt
caatataattt gtnaaaaactg agatacangt ttgacctata tctgcatttt gataattaaa
cnaatnnatt ctattnaat gttgtttcag agtcacagca cagactgaaa ctttttttga
atacctnaat atcacacttn tncttnnaat gatgttgaag acaatgtga catgccttna
gcatataatg tcgac 60
120
180
240
300
315

<210> 23

```

```

<211> 202
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(202)
<223> n = A,T,C or G

<400> 23
actaatccag tgggtgnaa ttccatttgtt ttgggcaact caggatatta aatttatnat      60
ttaaaaattc ccaagagaaa naaactccag gccctgattt gttcaactggg gaattttacc      120
aaatgttnca nnaaganatg acgctgattc tgtnaaatct ttttcagaag atagaggaga      180
acacccaccg nttcatttta tg                                202

<210> 24
<211> 365
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(365)
<223> n = A,T,C or G

<400> 24
ggattttcttg cccttttctc ccttttaag tatcaatgtt tgaaatccac ctgtaccacc      60
ctttctgcca tacaaccgct accacatctg gctccttagaa cctgtttgc tttcatagat      120
ggatctcgga accnagtgtt nacttcattt ttaaacccca ttttagcaga tngttgctn      180
tggctctgtct gtattcacca tggggctgt acacaccacg tgggttata gtcaaacaca      240
gtgcctcca ttgtggccac atgggagacc catnaccnna tactgcattt tggctgatn      300
acggcactgc atctnaccgg acntgggatt gaaccgggg tggcagcng aattgaacag      360
gatca                                365

<210> 25
<211> 359
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(359)
<223> n = A,T,C or G

<400> 25
gtttctgtt tcaacagtgc ttggacggaa cccggcgctc gttccccacc ccggccggcc      60
gcccatagcc agccctccgt cacctcttca ccgcaccctc ggactgcccc aaggcccccg      120
ccgcnctcc ngcgccncgc agccaccggc gcnccncca cctctcttgc tcccccgcnt      180
nacaacgcgt ccacacgtca ngttcgccng aactaccacc nggactcata ngccgcctc      240
aaccggccga tcaacacttgc gctctccccc ccgacntaa ccttccntg tcttacttac      300
nttaaccggc gnttatttttgc tttnaaaaga actttcccc aataacttgc ttcacccnt      359

<210> 26
<211> 400
<212> DNA

```

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(400)

<223> n = A,T,C or G

<400> 26

agtgaaacag tataatgtcaa aaggagtttgcannagcta cataaaaata ttagatatct	60
ttataatttc caataggata ctcatcaattt gacatattct agagaaaacca	120
ggtttctgg ttcagatttg aactctcaag agcttggaaat ttatcactcc catcctcaccg	180
acnacnaana aatctnaacn aacnagaanac caatgacttt tcttagatct gtcaaagaac	240
ttcagccacg agggaaaacta tcnccctnaa tactggggac tggaaagaga gggtagag	300
aatcacagtg aatcatagcc caagatcagc ttgcccggag cttaagctng tacgatnatt	360
acttacaggg accacttcac agtnngtnga tnaantgccn	400

<210> 27

<211> 366

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(366)

<223> n = A,T,C or G

<400> 27

gaatttctta gaaactgaag tttactctgt tccaagatatacttcaactg tcttaatcaa	60
aggcgctng aatcatagca aatattctca tctttcaact aacttaagt agttntcctg	120
gaattttaca tttccagaa aacactcctt tctgtatctg tgaaagaaag tgtgcctcag	180
gctgttagact gggctgcact ggacacctgc gggggactct ggctnagtgn ggacatggtc	240
agtatttattt ttcctcanac tcagcctgtg tagctntgaa agcatggaa acattttact	300
gcagtttacg tcatcccaca catcttggac tccnagaccc ggggaggtca catagtccgt	360
tatgna	366

<210> 28

<211> 402

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(402)

<223> n = A,T,C or G

<400> 28

agtgggagcc tcctccttcc ccactcagtt cttaatcc ccgaggcgca gctgggcnaa	60
ggaagtggcc agctgcagcg ctcctgcag gcagccaaacg ttcttgcctg tggcctgtgc	120
agacacatcc ttgccaccac cttaaccgtc catangcct gacacactgt gcacccactc	180
gctngctttt aagccccgtat nggctgcatt ctggggact tgacacaggc ncgtgatctt	240
gccagcctca ttgtccaccg tgaagagcat ggcaaaaagt ctgaggggag tgcatcttga	300
anagctcaa ggcttcattc agggccttng ctnaggcgcc nctctccatc tccnngaaata	360
acnagaggct ggtnngggtn actntcaata aactgcttcg tc	402

<210> 29

<211> 175
 <212> DNA
 <213> Homo sapien

<400> 29
 cggacgggca tgaccggtcc ggtcagctgg gtggccagtt tcagttcttc agcagaactg 60
 tctcccttct tgggggccga gggcttcctg gggaaagagga tgagtttggga gcggtactcc 120
 ttcagccgct gcacgttggt ctgcaggac tccgtggact tgttccgcct cctcg 175

<210> 30
 <211> 360
 <212> DNA
 <213> Homo sapien

<400> 30
 ttgtatttct tatgtatctct gatgggttct tctcgaaaat gccaaagtggaa agactttgtg 60
 gcatgcctca gatttaaatc cagctgaggc tccctttgtt ttcagttcca tgtaacaatc 120
 tggaaaggaaa cttcacggac aggaagactg ctggagaaga gaagcgtgtt agcccatttg 180
 aggtctgggg aatcatgtaa agggtaccca gacctcaattt ttagtttattt acatcaatga 240
 gttcttcag ggaacccaaac ccagaattcg gtgaaaaggc caaacatctt ggtgggattt 300
 gataaatgcc ttgggacctg gagtgctggg cttgtgcaca ggaagagcac cagccgctga 360

<210> 31
 <211> 380
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(380)
 <223> n = A,T,C or G

<400> 31
 acgctctaag cctgtccacg agctcaatacg ggaaggctgt gatgactaca gactttgcga 60
 acgctacgccc atggtttatg gatacaatgc tgccctataan cgctacttca ggaagcgcgg 120
 agggaccnaa tgagactgag ggaagaaaaaa aaatctctt ttttctggag gctggcacct 180
 gattttgtat cccccctgttnn cagcattncn gaaatacata ggcttataata caatgcttct 240
 ttccctgtata ttctcttggtc tggctgcacc ccttnttccc gccccagat tgataagtaa 300
 tgaaagtgcac ctgcagtnag ggtcaangga gactcancat atgtgattgt tccntnataa 360
 acttctggtg tgatactttc 380

<210> 32
 <211> 440
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(440)
 <223> n = A,T,C or G

<400> 32
 gtgtatggaa gcccctgact cctcacgtgc ctgatctgtg cccttggtcc caggtcaggc 60
 ccacccctg cacctccacc tgccccagcc cctgcctctg ccccaagtgg ggccagctgc 120
 cctcacttct ggggtggatg atgtgacccctt cctnggggaa ctgcggaaagg gacaagggtt 180

ccctgaagtc ttacggtcca acatcaggac caagtcccat ggacatgctg acagggtccc	240
caggggagac cgtntcanta gggatgtgtg cctggctgtg tacgtgggtg tgcaagtgcac	300
gtganaagca cgtggcggtct ctggggggcc atgtttgggg aaggaagtgt gcccnccacc	360
cttggagaac ctcaagtccccn gtagccccct gccctggcac agcngcatnc acttcaaggg	420
caccctttgg gggttgggggt	440
<210> 33	
<211> 345	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(345)	
<223> n = A,T,C or G	
<400> 33	
tatttaaca atgtttatta ttcatttatac cctctataga accaccaccc acaccgagga	60
gattatttgg agtgggtccc aacctagggc ctggactctg aaatctaact ccccacttcc	120
ctcattttgt gacttaggtg ggggcatgggt tcagtcagaa ctggtgtctc ctattggatc	180
gtgcagaagg aggacctagg cacacacata tggtgccac acccaggagg gttgattggc	240
aggctgaaag aaaaaagtct cccaaataaag gcactttac ctcaaagang gggtgggagt	300
tggtctgctg ggaatgttgt tggtgggggt gggagantt atttc	345
<210> 34	
<211> 440	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(440)	
<223> n = A,T,C or G	
<400> 34	
tgtaattttt ttattggaaa acaaataac aacttggaat ggattttgag gcaaattgtg	60
ccataaagcag atttaagtg gctaaacaaa gttaaaaag caagtaacaa taaaagaaaa	120
tgtttctggc acaggaccag cagtacaaaa aaatagtgtt cgagtacctg gataatacac	180
ccgttttgcg atagtgcac ttttaagtac atattgttga ctgtccatag tccacgcaga	240
gttacaactc cacacttcaa caacaacatg ctgacagtcc ctaaagaaaa ctactttaaa	300
aaaggcataa cccagatgtt ccctcatttgc accaactcca tctnagttt gatgtgcaga	360
agggcttana ttttcccaga gtaagccnca tgcaacatgt tacttgatca attttctaaa	420
ataagggtttt aggacaatga	440
<210> 35	
<211> 540	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(540)	
<223> n = A,T,C or G	
<400> 35	

atagatggaa	tttattaagg	ttttcacatg	tgatagcaca	tagtttaat	tgcatccaaa	60
gtactaaca	aaactctagc	aatcaagaat	ggcagcatgt	tatTTataa	caatcaacac	120
ctgtggctt	taaaatttgg	ttttcataag	ataattata	ctgaagtaaa	tctagccatg	180
ctttaaaaa	atgcttagg	tcactccaag	cttggcagtt	aacatttgc	ataaaacaata	240
ataaaacaat	cacaattaa	taaataacaa	atacaacatt	gtaggccata	atcatataca	300
gtataaggga	aaaggtggta	gtgttganta	agcagttatt	agaatagaat	accttggcct	360
ctatgcaa	atgtctagac	actttgatc	actcagccct	gacattcagt	tttcaaagtt	420
aggaaacagg	ttctacagta	tcatTTaca	gtttccaaca	cattgaaaac	aagtagaaaa	480
tgatgantt	atTTTatta	atgcattaca	tcctcaagan	ttatcaccaa	cccctcaggt	540
<210>	36					
<211>	555					
<212>	DNA					
<213>	Homo sapien					
<220>						
<221>	misc_feature					
<222>	(1)...(555)					
<223>	n = A,T,C or G					
<400>	36					
cttcgtgtgc	ttgaaaattt	gagcctgccc	ctcgccccat	aagcccttgt	tgggaactga	60
gaagtgtata	tggggcccaa	ntactgtgt	ccagaacaca	gagacagcag	cccantgcaa	120
tgctgtcgag	cattgcaaaac	gccatgtgt	gaactaggag	gaggaatatt	ccatcttggc	180
agaaaccaca	gcattggttt	ttttctactt	gtgtgtctgg	gggaatgaac	gcacagatct	240
gtttgactt	gttataaaaaa	tagggctccc	ccacctcccc	cntttctgt	tnctttattt	300
tagcantgt	gtctgcaagg	gagcccttan	cccctggcag	acananctgc	ttcagtgccc	360
ctttccctc	tgctaaatgg	atgttgatgc	actggaggtc	tttancctg	cccttgcatt	420
gnccctgctg	gaggaagana	aaactctgt	ggcatgaccc	acagtttctt	gactggangc	480
cntcaaccct	cttggttgaa	gccttgttct	gaccctgaca	tntgcttggg	cnctgggtng	540
gnctggcctt	ctnaa					555
<210>	37					
<211>	280					
<212>	DNA					
<213>	Homo sapien					
<220>						
<221>	misc_feature					
<222>	(1)...(280)					
<223>	n = A,T,C or G					
<400>	37					
ccaccgacta	taagaactat	gccctcggt	attcctgtac	ctgcatcatc	caacttttc	60
acgtggattt	tgcttggatc	ttggcaagaa	accctaattct	ccctccagaa	acagtggact	120
ctctaaaaaa	tatcctgact	tctaataaca	ttgatntcaa	gaaaatgacg	gtcacagacc	180
aggtgaactg	ccccnagctc	tcgtaaccag	gttctacagg	gaggctgcac	ccactccatg	240
ttncttctgc	ttcgctttcc	cctacccac	cccccgccat			280
<210>	38					
<211>	303					
<212>	DNA					
<213>	Homo sapien					
<220>						

```

<221> misc_feature
<222> (1)...(303)
<223> n = A,T,C or G

<400> 38
catcgagctg gttgtcttct tgcctgccct gtgtcgtaaa atgggggtcc cttactgcat      60
tatcaaggga aaggcaagac tgggacgtct agtccacagg aagacctgca ccactgtcgc      120
cttcacacag gtgaactcgg aagacaaagg cgcttggct nagctggtn aagctatcag      180
gaccaattac aatgacngat acgatnagat ccgcncntcac tggggtagca atgtccctggg      240
tcctaagtct gtggctcgta tcgcccnaagct cgaanaggcn aangctaaag aacttgccac      300
taa                                         303

<210> 39
<211> 300
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(300)
<223> n = A,T,C or G

<400> 39
gactcagcgg ctgggtgctct tcctgtgcac aagcccagca ctccagggtcc caaggcattt      60
atcaaatccc accaagatnt ttggcttttgc acccgaattt tgggttttgt tccctnaaag      120
aactcattga tgtaaatnac tnaaaatgtgag gtctgggtac cctttacatg attccccaga      180
cctcanatgg gctaacacgc ttctcttctc cagcagttt cctntccgtg aagtacctt      240
ccagattgtt acatgaaact gaanacaaag ggagcctcag ctngattaa atctggagca      300

<210> 40
<211> 318
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(318)
<223> n = A,T,C or G

<400> 40
cccaacacaa tggctgagga caaatcagtt ctctgtgacc agacatgaga aggttgccaa      60
tgggctgttg ggcgaccaag gccttcccg agtcttcgtc ctctatgagc tctcgcccat      120
gatggtaag ctgacggaga agcacaggc cttcacccac ttccctgacag gtgtgtgcgc      180
catcattggg ggcatgttca cagtggttgg actcatcgat tcgctcatct accactcagc      240
acgagccatc cagaaaaaaaaa ttgatctngg gaagacnacg tagtcacccct cggtncttcc      300
tctgtctcct cttctcc                                         318

<210> 41
<211> 302
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(302)

```

<223> n = A,T,C or G

<400> 41

acttagatgg ggtccgttca ggggataacca gcgttcacat ttttcctttt aagaaagggt 60
cttggcctga atgtccccca tccggacaca ggctgcattgt ctctgnagt gtcaaagctg 120
ccatnaccat ctcggtaacc tactcttact ccacaatgtc tatnttcaact gcagggctct 180
ataatnagtc cataatgtaa atgcctggcc caagacntat ggcctgagtt tatccnaggc 240
ccaaacnatt accagacatt cctttanat tgaaaacgga tntcttccc ttggcaaaga 300
tc 302

<210> 42
<211> 299
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(299)
<223> n = A,T,C or G

<400> 42

cttaataagt ttaaggccaa ggcccgttcc atttcttctag caactgacgt tgccagccga 60
ggtttgaca tacctcatgt aaatgtggtt gtcaactttg acattcctac ccattccaag 120
gattacatcc atcgagtagg tcgaacagct agagctggc gctccggaaa ggcttattact 180
tttgtcacac agtatgtatgtt ggaacttcc cagcgcatag aacacttnat tggaaagaaa 240
ctaccagggtt ttccaacaca ggatgatgag gttatgatgc tnacggaacg cgtcgctna 299

<210> 43
<211> 305
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(305)
<223> n = A,T,C or G

<400> 43

ccaaacaatgt caagacagcc gtctgtgaca tcccacctcg tggcctcaan atggcagtca 60
ccttcattgg caatagcaca gcncntccggg agctttcaa ggcacatctcg gagcagttca 120
ctgccatgtt ccggccggaaag gccttcctcc actggtagac acggcagggc atggacaaga 180
tggagttcac cgaggctgag agcaacatga acgacctcgct ctctnagtat cagcagtacc 240
gggatgccac cgccagaaana ggaggaggat ttccggtnagg aggccgaaga aggaggcctg 300
aggca 305

<210> 44
<211> 399
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(399)
<223> n = A,T,C or G

```

<400> 44
tttctgtggg gaaaacctga tctcgacnaa attagagaat tttgtcagcg gtatttcggc      60
tggAACAGAA cgaaaaacnga tnaatctctg tttccctgtat taaagcaact cgatncccag      120
cagacacagc tccnaattga ttcccttctt ngatttagcac aacagggaga aagaanatgc      180
ttaacgtatt aagagccnga gactaaacag agcttgaca tgtatgctt ggaaagagaa      240
agaaggcagcn gcccgcncaa ttngaagcng tttctgttgc cntgganaaa gaatttgagc      300
ttcttattt gccaacgaa aaaccccgaa ananaggcnt tacnataacct tngaaaantc      360
tccngccnna aaaagaaaaga agctttcnga ttcttaacc      399

<210> 45
<211> 440
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(440)
<223> n = A,T,C or G

<400> 45
gcgggaggcag aagcttaaagc caaagcccaa gagagtggca gtgccagcac tgggccagt      60
accagtagcca ataacagtgc cagtgcaggc gccagcacca gtgggtggctt cagtgctgg      120
gccagcctga ccgcccactt cacattttgg ctttcgctg gccttgggtt agctgggtgcc      180
agcaccaggc gcagctctgg tgcctgttgtt ttccctaca agtgagattt taggtatctg      240
ccttggtttc agtggggaca tctggggctt anggggcngg gataaggagc tggatgattc      300
taggaaggcc cangttggag aangatgtgn anagtgtgcc aagacactgc ttttggcatt      360
ttattccctt ctgtttgctg gangtcaatt gacccttnna ntttcttta cttgtgtttt      420
canatatngt taatcctgcc      440

<210> 46
<211> 472
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(472)
<223> n = A,T,C or G

<400> 46
gctctgtaat ttacatccccaa acacccccc ttgacctcac attcccttcc ggccacctct      60
gtttctctgt tcctcttcac agaaaaact gttcaaaaaga gttgttgatt acttcattt      120
ccactttctc acccccatc tcccccaat taactctctt tcattcccat gatgccatta      180
tgtggctntt attanagtca ccaacccat tctccaaaac anaagcaaca aggactttga      240
cttctcagca gcactcagct ctggtnctt aaacacccccc gttacttgctt attccctcta      300
cctcataaca atcccttcc cagcctctac tgctgccttc tctgagttct tcccagggtc      360
ctaggctcag atgtatgtta gctcaaccct gctacacaaa gnaatctctt gaaaggctgt      420
aaaaatgtcc atncntgtcc tggatgtat ctnccangna naataacaaa tt      472

<210> 47
<211> 550
<212> DNA
<213> Homo sapien

<220>

```

```

<221> misc_feature
<222> (1)...(550)
<223> n = A,T,C or G

<400> 47
ccttcctccg cctggccatc cccagcatgc tcatgctgtg catggagtgg tgggcctatg      60
aggcggag ctccctcagt ggtctgtatg aggatggatg acggggactg gtgggaacct      120
gggggcctcg tctgggtgca aggcacacgc tgtctttctt caccaggcat cctccggcatg      180
gtggagctgg gcgcctcagtc catcgtgtat gaactggcca tcattgtgtat catggccct      240
gcaggcttca gtgtggctgc cagtgtccgg gtangaaacg ctctgggtgc tggagacatg      300
gaagcaggca cgaaagtccct ctaccgttgc cctgctgatt acagtgtct ttgctgtanc      360
cttcagtgta ctgctgttaa gctgtaaagg tcacntgggg tacattttta ctaccgaccg      420
agaacatcat taatctgggtg gtcaggtgg ttccaaattta tgctgtttcc cacctctttg      480
aagctttgc tgctcaggtt caccgcattt ttgaaaagta aacaacgtgc ctcggagtgg      540
gaattctgct      550

<210> 48
<211> 214
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(214)
<223> n = A,T,C or G

<400> 48
agaaggacat aaacaagctg aacctgcccc aagacgtgtga tatcagcttc tcagatccag      60
acaacccct caacttcaag ctggcatct gtcctgatna gggcttctac nagagtggga      120
agtttgtgtt cagtttaag gtggccagg gttacccgca tgatcccccc aaggtgaagt      180
gtgagacnat ggtctatcac cccnacattt acct      214

<210> 49
<211> 267
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(267)
<223> n = A,T,C or G

<400> 49
atctgcctaa aatttattca aataatgaaa atnaatctgt tttaaagaaat tcagtcattt      60
agtttttagg acaactatgc acaaatgtac gatggagaat tcttttggta tnaactctag      120
gtngagaaac ttaatccaac cggagcttt gtgaaggtca gaanacagga gaggaatct      180
tggcaaggaa tggagacnnga gtttgcattt tgcaagctaga gttaatngtt nttaatggga      240
ctgctttgt gtctcccang gaaagtt      267

<210> 50
<211> 300
<212> DNA
<213> Homo sapien

<220>

```

```

<221> misc_feature
<222> (1)...(300)
<223> n = A,T,C or G

<400> 50
gactgggtca aagctgcattt aaaccaggcc ctggcagcaa cctggaaatg gctggagggtg      60
ggagagaacc tgacttctct ttccctctcc ctccctccaa attactggaa ctctgtcctg      120
ttgggatctt ctgagcttgc ttccctgctg ggtggacag aggacaaaagg agaaggggagg      180
gtctagaaga ggcagccctt ctttgtcctc tgggttaat gagcttgacc tanagtagat      240
ggagagacca anagcctctg attttaatt tccataanat gtcnaagta tatntntacc      300

<210> 51
<211> 300
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(300)
<223> n = A,T,C or G

<400> 51
gggtaaaaatc ctgcagcacc cactctggaa aatactgctc ttaattttcc tgaagggtggc      60
cccctatttc tagttggtcc aggattaggg atgtggggta tagggcattt aaatcccttc      120
aagcgctctc caagcaccccc cggcctgggg gtnagttct catcccgcta ctgctgctgg      180
gatcaggtn aataaatgga actttcttg tctggcctcc aaagcagcct aaaaactgag      240
gggctctgtt agaggggacc tccaccctnn ggaagtccga ggggctnggg aagggtttct      300

<210> 52
<211> 267
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(267)
<223> n = A,T,C or G

<400> 52
aaaatcaact tcntgcatta atanacanat tctanancag gaagtgaana taattttctg      60
cacctatcaa ggaacnnact tgattgcctc tattnaacan atatatcgag ttnctatact      120
tacctgaata ccnccgcata actctcaacc nanatncntc nccatgacac tcnttcttna      180
atgctantcc cgaattcttc attatatcng tcatgttcgn cctgntnata tatcagcaag      240
gtatgtncnn taactgccga nncaang                                267

<210> 53
<211> 401
<212> DNA
<213> Homo sapien

<400> 53
agscttttagc atcatgtaga agcaaactgc acctatggct gagataggtg caatgaccta      60
caagattttg tgtttcttag ctgtccagga aaagccatct tcagtcttgc tgacagtcaa      120
agagcaagtg aaaccatttc cagcctaaac tacataaaaag cagccgaacc aatgattaaa      180
gacctcttaag gctccataat catcattaaa tatgcccaaa ctcattgtga ctttttattt      240

```

tatatacagg attaaaatca acattaaatc atcttattta catggccatc ggtgctgaaa	300
ttgagcattt taaatagtac agtaggctgg tatacattag gaaatggact gcactggagg	360
caaataaaaa actaaagaaa tttagataggc tggaaatgct t	401
<210> 54	
<211> 401	
<212> DNA	
<213> Homo sapien	
<400> 54	
cccaacacaa tggataaaaa cacttatagt aaatgggac attcaactata atgatctaag	60
aagctacaga ttgtcatagt tggttcctg cttacaaaa ttgctccaga tctgaatgc	120
cagtttgacc tttgtcttct ataataatttc cttttttcc cctcttgaa tctctgtata	180
tttgattctt aactaaaatt gttctcttaa atattctgaa tcctggtaat taaaagttt	240
ggtgtatttt ct当地tcc aaggaaagaa ctactagcta caaaaaatata tttgaaataa	300
gcattgtttt ggataaggt acatattttg gttgaagaca ccagactgaa gtaaacagct	360
gtgcacccaa tttattatag tttgttaagt aacaatatgt a	401
<210> 55	
<211> 933	
<212> DNA	
<213> Homo sapien	
<400> 55	
tttactgctt ggcaaaagtac cctgaggcatc agcagagatg ccgagatgaa atcagggAAC	60
tccttagggga tgggtcttctt attacctggg aacaccttag ccagatgcct tacaccacga	120
tgtgcataaa ggaatgcctc cgccctctacg caccggtagt aaactatccc ggtaactcg	180
caaaccatc acctttccag atggacgctc cttacctgca ggaataactg tgtttatcaa	240
tatTTGGGCT CTTCAACCACCA ACCCCTATTCTGGGAAGAC CCTCAGGTCT TTAACCCCTT	300
gagattctcc agggaaaatt ctgaaaaat acatccctat gccttcatac catttcagc	360
tggattaagg aactgcattt ggcagcattt tgccataatt gagtgtaaag tggcagtggc	420
attaactctg ctccgcttca agctggctcc agaccactca aggccaccca gctgtcgta	480
agttgcctca agtccaagaa tggaaatccat gtgttgcaa aaaaagttt ctaattttaa	540
gtcctttcg tataagaatt aakgagacaa tttcctacc aaaggaaagaa caaaaggata	600
aatataatac aaaatataatg tatatggttt ttgacaaat tatataactt aggataacttc	660
tgactggttt tgacatccat taacagtaat ttaattttct ttgctgtatc tggtaaaacc	720
cacaaaaaca cctgaaaaaaaaa ctcaagctga gttcaatgc gaaggaaat gattggttt	780
ggttaactagt ggttagagtgg ctttcaagca tagtttgatc aaaactccac tcagtatctg	840
cattactttt atctctgcaa atatctgcat gatagtttta ttctcagttt tctttccccca	900
taataaaaaa tatctgccaa aaaaaaaaaaaa aaa	933
<210> 56	
<211> 480	
<212> DNA	
<213> Homo sapien	
<400> 56	
ggctttaag cattttgtc tttgtctccctt gatttcagg tcaccaccat gaaggttctta	60
gcagtcctgg tacttctggg agtttccatc ttctggctt ctgcccagaa tccgacaaca	120
gctgctccag ctgacacgtt tccagctact ggtcctgctt atgatgaagc ccctgtatgt	180
gaaaccactg ctgctgcaac cactgctgctt ctaccactgc aaccaccgtt	240
gcttctacca ctgctcgtaa agacattcca gtttaccca aatgggttgg ggatctcccc	300
aatggtagag tttgtccctg agatgaaatc agtttgcattt ggtcacaact	360
attcatgctt cttgttgatctt catccaacta cttaccttgc ctacgatatac cccttttatct	420
ctaatcagtt tattttctttt caaataaaaaa ataaactatga gcaacaaaaa aaaaaaaaaa	480

<210> 57		
<211> 798		
<212> DNA		
<213> Homo sapien		
<400> 57		
agcctacctg gaaagccaac cagtccatcat aatggacaag atccaccagc tcctcctgtg	60	
gactaacttt gtgatatggg aagtgaaaat agttaacacc ttgcacgacc aaacgaacga	120	
agatgaccag agtactctta accccttaga actgttttc ctttgtatc tgcaatatgg	180	
gatggatttg tttcatgag cttctagaaa ttcaacttgc aagtttattt ttgcttcctg	240	
tgttactgcc attcctatTT acagtatatt tgagtgaatg attatatttt taaaaagttt	300	
catggggctt tttgggttgc cctaaactta caaacattcc actcattctg tttgttaactg	360	
tgattataat tttgtgtata atttctggcc tgattgaagg aaatttgaga ggtctgcatt	420	
tatataatTTT aaatagattt gataggTTT taaattgttt ttttcataaa ggtatTTATA	480	
aagttatTTG gggTTgtctg ggattgtgtg aaagaaaatt agaacccccgc ttttatttaca	540	
tttaccttgg tagtttattt gtggatggca gttttctgtt gttttggggta ctgtggtagc	600	
tcttggattt ttttgcAAAT tacagctgaa atctgtgtca tggattaaac tggcttatgt	660	
ggctagaata ggaagagaga aaaaatgaaa tggTTgttta ctaatTTTactccattt	720	
aaaatttttta atgttaagaa aaccttaaat aaacatgatt gatcaatatg gaaaaaaaaaa	780	
aaaaaaaaaaa aaaaaaaaaa	798	
<210> 58		
<211> 280		
<212> DNA		
<213> Homo sapien		
<400> 58		
ggggcagctc ctgaccctcc acagccaccc ggtcagccac cagctggggc aacgagggtg	60	
gaggtcccac tgagcctctc gcctgcccccc gccactcgcc tggtgcttgc tgatccaagt	120	
cccccgcctg gtccccccaca aggactccca tccaggcccc ccctgcctg cccctgtca	180	
tggaccatgg tcgtgaggaa gggctcatgc cccttattta tggaaaccat ttcattctaa	240	
cagaataaac cgagaaggaa accagaaaaaaa aaaaaaaaaaaa	280	
<210> 59		
<211> 382		
<212> DNA		
<213> Homo sapien		
<400> 59		
aggcggggagc agaagctaaa gccaaagccc aagagagtgg cagtccacactggccca	60	
gtaccagtac caataaacagt gccagtgcac gtccacgcac cagtgggtggc ttcaagtgc	120	
gtgccagcct gaccgcact ctcacatttgc ggcttgcgc tggccttgc ggagctgggt	180	
ccagcaccag tggcagctt ggtgcctgtt gttcttccta caagtgagat ttttagatatt	240	
gttaatctcg ccagtcttc tcttcagacc agggtgcatc ctcagaaacc tactcaacac	300	
agcactcttag gcagccacta tcaatcaattt gaagttgaca ctctgcattt aatctatttgc	360	
ccattaaaaaa aaaaaaaaaaaa aa	382	
<210> 60		
<211> 602		
<212> DNA		
<213> Homo sapien		
<400> 60		
tgaagagcccg cgcgggtggag ctgctgcccc atggactgc caaccttgcc aagctgcagc	60	

ttgtgggtgga gaatagtgc cagcgggtca tccacttggc gggtcagtgg gagaagcacc 120
gggtccccatc ctcgtgagta ccgcccactcc gaaagctgca ggattgcaga gagctgaat 180
cttctcgacg gctggcagag atccaagaac tgaccaggag tgtccggcgc gctgctgaag 240
aggccccgca gaaaggaggag gtctataaagc agctgatgtc agagctggag actctgccc 300
gagatgtgtc ccggctggcc tacaccccagc gcacccctggaa gatcggtggc aacatccgga 360
agcagaagga agagatcacc aagatcttgt ctgatacgaa ggagcttcag aaggaaatca 420
actccctatc tgggaagctg gaccggacgt ttgcggtgac tgatgagctt gtgttcaagg 480
atgccaagaa ggacgatgct gttcggaaagg cctataagta tctagctgct ctgcacgaga 540
actgcagcca gctcatccag accatcgagg acacaggcac catcatgcgg gaggttcgag 600
ac 602

```
<210> 61  
<211> 1368  
<212> DNA  
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(1368)
<223> n = A,T,C or G
```

<400> 61

ccagtggcgc cgcgtaatac gactcaactat agggcgaatt gggtaccggg cccccctcg
agcggccgccc cttttttttt ttttttttatt gatcagaatt caggctttat tattgagcaa
tgaaaacagc taaaactaa ttccaagcat gtgtagttaa agtttgc当地 gtgggatatt
gttcacaaaa cacattcaat gtttaaacac tatttattt aagaacaaaa tatattaaa
atttgttgc tctaaaaagc ccatttccct ccaagtctaa actttgtaat ttgatattaa
gcaatgaagt tattttgtac aatcttagtta aacaagcaga atagcactag gcagaataaa
aaattgcaca gacgtatgc attttccaag atagcattct ttaaattcag tttcagctt
ccaaagattt gttgccata atagacttta acatataatg atggctaaaa aaaataagta
tacaaaaatg taaaaaaagga aatgtaaatgc cactctcaat ctcataaaag gtgagagtaa
ggatgctaaa gcaaaataaa tgtagggttct tttttctgt ttccgcttat catgcaatct
gcttccttga tatgccttag ggttaccat ttaagttaga gttgtatg caatggtggg
aatgaaaatt gatcaaataat acacccctgtc atttcatttc aaattgcggg ctggaaactt
ccaaaaaaaaag ggtaggcatg aaaaaaaaaa aaatcmatac agaacctttt caggggtttt
kgktctgata tggcagacar gatacaagtcc acccaggag atggagcaat tcaaaaataag
ggtaatggc tgacaaggta ttattgcccag catgggacag aatgagcaac aggctaaaaa
gttttgat tatatacgac cttagtctc tgatgttaggg aattttgtt agtcaacat
acgctaaact tccaaaggaa aatctttcag gtggcttaag ctgcctttt tagagtgtat
agttgcattt ctactgtat tttttgaaaa caaactgggt ttgtacaagt gagaagact
agagagaaaat ttttagtct gtttagcaga agccattttt aatgcgtgca catggatcaa
tatttctgat cccctataacc ccaggaagggg caaaatccca aagaaatgtt ttagcaaaat
tggctgatgc tatcatattt gatggacat tgatcttgc caacacaatg gaattccacc
acactggact agtggatcca ctgttcttag agcggccggc caccgcgtg gagctccagc
ttttgttccc tttagtgagg gtttaattgcg cgcttggcgt aatcatnn

<210> 62
<211> 924
<212> DNA
<213> *Homo sapien*

```
<220>
<221> misc_feature
<222> (1)...(924)
<223> n = A,T,C or G
```

<400> 62

caaaggaca ggaacagctt gnaaaagtact gnacatncctn cctgcagggga ccagcccttt	60
gcctccaaaa gcaataggaa atttaaaaaga tttncaactga gaaggggncc acgtttnart	120
tntnaatgtn tcargnanar tnccttncaa atgncrnctn cactnactnr gnattgggt	180
tnccgnrtnm mgnactatnt caggtttcaa aaactggatc tgccacttat cagttatgtg	240
accttaaaga actccgttaa tttctcagag cctcagttc cttgtctata agttggaggt	300
aatattaata ctatcatttt tccaaggatt gatgtgaaca ttaatgaggt gaaatgacag	360
atgtgtatca tggttcccaa taaacatcca aaatatagta cttactattt tcattattat	420
tacttggttt aagctaaaga cctcacaata gaatcccattc cagcccacca gacagagyc	480
tgagtttctt agtttggaaag agctttaaaa taacaacktc tagtgtcaat tctatacttg	540
ttatggtcaa gtaactgggc tcagcattttt acatttcatttgc tctcttttaag ttcttagcaat	600
gtgaaggcagg aactatgatt atattgacta cataaatgaa gaaattggagg ctcagataca	660
ttaagtaatt ctcccagggt cacacagcta gaactggcaaa agcctggat tgatccatga	720
tcttccagca ttgaagaatc ataaatgtaa ataactgca ggcctttcc tcagaagagc	780
tcctggtgct tgcaccaacc cactagcact tggctctac aggggaacat ctgtgggcct	840
gggaatcact gcacgtcgca agagatgtt cttctgatga attattgttc ctgtcagtgg	900
tgtgaaggca aaaaaaaaaaaa aaaa	924

<210> 63

<211> 1079

<212> DNA

<213> Homo sapien

<400> 63

agtccccaga actcaataat ctcttatgtt ttcttttggaa gacttatttt aaatattaac	60
tatccggtg cctgaatggaa aaaatataaa cattagctca gagacaatgg ggtacctgtt	120
tggaaatccag ctggcagctt taagcaccgt tgaaaactct gacaggctt gtgcctttt	180
tattaaatgg cctcacatcc tgaatgcagg aatgtgttcg tttaaataaaa cattaatctt	240
taatgtgaa ttctgaaaac acaaccataa atcatagttt gttttctgt gacaatgatc	300
tagtacatta ttccctccac agcaaaccata cctttccaga aggtggaaat tgtatggca	360
acaatcaggg caaaaacccac acttggaaaag catttacaa tattatatct aagtgcaca	420
gaagacccca gtgatcaacta ggaaatctac cacagtccag tttttctaat ccaagaaggt	480
caaacttcgg ggaataatgt gtccctcttc tgctgctgt ctggaaaata ttcgatcaaa	540
acgaagtttta caagcagcag ttattccaaag attagagttc atttgtgtat cccatgtata	600
ctggcaatgtt ttaggtttgc ccaaaaactc ccagacatcc acaatgttgc tggtaaacc	660
accacatctg gtaacctctc gateccttag atttgcattt cctgcaaaata taactgttagc	720
tgactctggaa gcctcttgca ttttctttaa aaccattttt aactgattca ttcgttccgc	780
agcatgccct ctggtgctt ccaaatggga tgcataagg caaagctcat ttcctgacac	840
attcacatgc acacataaaa gtttctcat cattttggta cttggaaaag gaataatctc	900
ttggctttt aatttcactc ttgatttttca acacattata gctgtgaaat atccttcttc	960
atgacctgta ataatctcat aattacttga tctcttctt aggtagctat aatatggggg	1020
aataacttcc tgcataaaaata tcacatctgg gctgtacaaa gctaagttagg aacacaccc	1079

<210> 64

<211> 1001

<212> DNA

<213> Homo sapien

<400> 64

gaatgtgcaa cgatcaagtc agggtatctg tggatccac cactttgagc atttacgtat	60
tctataatgtc aggaacattt caagttatct gttcttagcaaa ggaaatataaa aataacttata	120
gttaactatg gcctatctac agtgcaacta aaaacttagat ttatccctt tccacctgtg	180
ggtttggatt catttaccac cctctttca ttccctttctt caccacacaca ctgtgcccggg	240
cctcaggcat atactatttctt actgtctgtc tctgtaaaggaa ttatcatttt agcttccacca	300

tatgagagaa	tgcatgcaaa	gtttttctt	ccatgtctgg	cttatttcac	ttaacataat	360
gacctccgt	tccatccatg	ttatTTtat	tacccaatag	tgttcataaaa	tatatacaca	420
cacatatata	ccacattgca	tttgtccaaat	tattcattga	cggaaactgg	ttaatgttat	480
atcgTTgcta	ttgtggatag	tgctgcaata	aacacgcaag	ttggggatata	atTTgaagag	540
ttttttgtt	gatgttcctc	caaattttaa	gattttttt	tctatgtttt	tgaaaatggc	600
gttagtattt	tcatagagat	tgcattgaat	ctgttagattt	ctttgggtaa	gtatggttat	660
tttgatggta	ttaatTTTT	cattccatga	agatgagatg	tcttccatt	gtttgtgtcc	720
tctacatTTT	ctttcatcaa	agttttgtt	tatTTTgaa	gtagatgtat	ttcaccttat	780
agatcaagtg	tattccctaa	atatTTTatt	ttttagtcta	ttttagatga	aattgccttc	840
ttgattttt	tttcaCTTAA	ttcattattta	gtgtatggaa	atgttatggaa	tttttatttg	900
ttggTTTTTA	atcaaaaaact	gtattaaact	tagatTTTT	tgtggagtt	ttaagTTTT	960
ctagataaa	gatcatgaca	tctaccaaaa	aaaaaaaaaa	a		1000

```
<210> 65  
<211> 575  
<212> DNA  
<213> Homo sapien
```

```

<400> 65
acttgatata aaaaggatat ccataatgaa tattttatac tgcatccctt acatttagcca 60
ctaaatacgt tattgcttga tgaagacctt tcacagaatc ctatggattg cagcattca 120
cttggctact tcataacccat gccttaaaga ggggcagttt ctcaaaaagca gaaacatgcc 180
gccagttctc aagtttcct cctaactcca tttgaatgta agggcagctg gcccccaatg 240
tggggaggtc cgaacatccc ctgaattccc attttcttgt tcgcggctaa atgacagttt 300
ctgtcattac ttagattccc gatcttccc aaagggtttg atttacaaag aggccagcta 360
atagccagaa atcatgaccc tgaaagagag atgaaatttc aagctgttag ccagggcagga 420
gctccagttt ggcaaaagggtt cttgagaatc agccatttgg tacaaaaaaag atttttaaag 480
cttttatgtt ataccatggaa gccatagaaaa ggctatggat tgtttaagaa ctatTTaaa 540
qtqttccaaqa cccaaaaaaqq aaaaaaaaaa aaaaaa 575

```

```
<210> 66  
<211> 831  
<212> DNA  
<213> Homo sapien
```

<400> 66	
atggggctcc ttctgtctaaa cagccacatt gaaatggttt aaaagcaagt cagatcagg	60
gatttgtaaa attgtattta tctgtacatg tatgggcttt taattcccac caagaaagag	120
agaaaattatc ttttagtta aaaccaaatt tcactttca aaatatcttc caacttattt	180
atgggtgtc actcaattgc ctatataat atatataat gtgtgtgtgt gtgtgtgcgc	240
gtgagcgcac gtgtgtgtat gcgtgcgcgcgtat gtgtgtgtat gtgtattatc agacataggt	300
ttctaacttt tagatagaag aggagcaaca tctatgccaa atactgtgca ttctacaatg	360
gtgctaattct cagacctaaa tgatactcca ttaatttaaa aaaagagtt taaataatta	420
tctatgtgcc tgtatttccc ttttgagtgc tgcacaacat gttaacatata tagtgtaaaa	480
gcagatgaaa caaccacgtg ttctaaagtc tagggattgt gctataatcc ctatttagtt	540
caaaaattaac cagaattctt ccatgtgaaa tggacccaaac tcatattatt gttatgtaaa	600
tacagagttt taatgcagta tgacatccccca cagggaaaaaa gaatgtctgt agtgggtgac	660
tgttatcaaa tattttatag aataacaatga acggtaaca gactggtaac ttgtttgagt	720
tcccatgaca gatttgagac ttgtcaatag caaatcattt ttgtatttaa atttttgtac	780
tqattgtaaa aacatcattta aatatcttta aaqtaaaaaa aaaaaaaaaa a	831

<210> 67
<211> 590
<212> DNA
<213> *Homo sapien*

<400> 67
gtgctctgtg tattttta ctgcattaga cattgaatag taatttgcgt taagatacgc 60
ttaaaggctc ttgtgacca tgttccctt tgttagcaata aaatgttt tacgaaaact 120
ttctccctgg attagcagt taaatgaaac agagtcac aatgaaatga gtatttaaaa 180
taaaaatttg cctaattgtt tcagttcagc tcacaaggat ttaagatga ttgagaagac 240
ttgaattaaa gaaaaaaaaa ttctcaatca tattttaaa atataagact aaaattgtt 300
ttaaaacaca ttccaaatag aagtgagtt gaactgacct tatttatact ctttttaagt 360
ttgttcctt ccctgtgcc tgtgtcaat cttcaagtct tgctgaaaat acatttgata 420
caaagtttc tgttagttgtg tttagttttt tgcattgtct gttttggct gaagaaccaa 480
gaaggcagact ttcttttaa aagaattatt tctcttcaa atatttctat ctttttaaaa 540
aaattcctt ttatggcttatacctaca tatttaaaaa aaaaaaaaaa 590

<210> 68
<211> 291
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(291)
<223> n = A,T,C or G

<400> 68
gttccctttt ccggtcggcg tggcttgcg agtggagtgt ccgctgtgcc cgggcctgca 60
ccatgagcgt cccggcccttc atcgacatca gtgaagaaga tcaggctgtc gagttcgtg 120
cttatctgaa atctaaagga gctgagattt cagaagagaa ctcggaaaggt ggacttcatg 180
ttgatttagc tcaaattattt gaagcctgtg atgtgtgtct gaaggagatg gataaagatg 240
ttgaaagtgt gatgaacagt ggggnatct actcttgatc cggaancnac 291

<210> 69
<211> 301
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G

<400> 69
tctatgagca tgccaaggct ctgtggagg atgaaggagt gcgtgcctgc tacgaacgct 60
ccaaacgagta ccagctgatt gactgtgccccc agtacttctt ggacaagatc gacgtgatca 120
agcaggctga ctatgtgccg agcgatcagg acctgctcg ctggcgtgtc ctgacttctg 180
gaatcttga gaccaagttc caggtggacn aagtcaactt ccacatgntt gacgtgggtg 240
gccagcgcga tgaacgcccgc aagtggatcc agtgcattcaa cgatgtgact gccatcatct 300
t

<210> 70
<211> 201
<212> DNA
<213> Homo sapien

<400> 70
gcggctttc ctcggcagc ggaagcggcg cggcggtcg agaagtggcc taaaacttcg 60

gcgttgggtg aaagaaaatg gcccgaacca agcagactgc tcgtaagtcc accgggtggga	120
aagcccccccg caaacagctg gccacgaaag ccgccaggaa aagcgctccc tctaccggcg	180
gggtgaagaa gcctcatcgc t	201
<210> 71	
<211> 301	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(301)	
<223> n = A,T,C or G	
<400> 71	
gccggggtag tcgcccncgc cgccgcccgt gcagccactg cagggcacccgc tgccgcccgc	60
ttagtagtgg gcttaggaag gaagaggtca tctcgctcg agcttcgctc ggaagggtct	120
ttgttcctcg cagccctccc acggaaatga caatggataa aagttagctg gtacanaaaag	180
ccaaactcgc tgagcaggct gagcgatatg atgatatggc tgcagccatg aaggcagtca	240
cagaacaggg gcatgaactc ttcaacgaag agagaaatct gctctctggt gcctacaaga	300
a	301
<210> 72	
<211> 251	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(251)	
<223> n = A,T,C or G	
<400> 72	
cttgggggggt gttgggggag agactgtggg cctggaaata aaacttgtct cctctaccac	60
caccctgtac cctagcctgc acctgtccac atctctgcaa agttcagctt cttcccccag	120
gtctctgtgc actctgtctt ggatgctctg gggagctcat gggtgagggaa gtctccacca	180
gagggaggct caggggactg gttggccag ggatgaatat ttgagggata aaaattgtgt	240
aagagccaan g	251
<210> 73	
<211> 895	
<212> DNA	
<213> Homo sapien	
<400> 73	
tttttttttt ttttcccag gccctctttt tatttacagt gataccaaac catccacttg	60
caaattttttt ggtctccat cagctggaat taagtaggtt ctgtgtatct ttgagatcat	120
gtatttgtct ccactttgggt ggatacaaga aagaaggca cgaacagctg aaaaagaagg	180
gtatcacacc gctccagctg gaatccagca ggaacctctg agcacccac agctgaacac	240
ttaaaagagg aaagaaggac agctgctttt catttattttt gaaagcaaat tcatttggaaa	300
gtgcataaat ggtcatcata agtcaaactgt atcaattaga ccttcaacct aggaaacaaa	360
attttttttt tctatttaat aatacaccac actgaaatta tttgccaatg aatccaaag	420
atttggtaca aatagtacaa ttctgttatttgc ttccctctt tcctttcttc agacaaacac	480
caaataaaaat gcaggtgaaa gagatgaacc acgactagag gctgacttag aattttatgc	540
tgactcgtatc taaaaaaaaat tatgttgggtt aatgttaatc tatctaaaaat agacgtttt	600

gggaatgctt ttcaaagaag gtcaagtaac agtcatacag ctagaaaagt ccctgaaaaa	660
aagaatgtt aagaagtata ataacccttt caaaaccac aatgcagctt agtttcctt	720
tatTTATTTG tggTCatgaa gactatcccc atttctccat aaaatcctcc ctccatactg	780
ctgcattatg gcacaaaaga ctctaagtgc caccagacag aaggaccaga gtttctgatt	840
ataaacaatg atgctggta atgtttaat gagaacattg gatatggatg gtcag	895
<210> 74	
<211> 351	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(351)	
<223> n = A,T,C or G	
<400> 74	
tgtgcncagg ggatgggtgg gcngtggaga ngatgacaga aaggctgqaa ggaanggggg	60
tgggTTTgaa ggccangggc aagggggnct caggtccgnt tctgnnaagg gacagccttg	120
aggaaggagn catggcaagc catagctagg ccaccaatca gattaagaaa nnctgagaaa	180
nctagctgac catcaactgtt ggtgnccagt ttcccaacac aatggaatnc caccacactg	240
gactagngga nccactagtt ctagagcggc cgccaccgcg gtggAACCCC aactttgcc	300
cctttagnga gggtaatttgcgcttgc ntaatcatgg tcataagctg t	351
<210> 75	
<211> 251	
<212> DNA	
<213> Homo sapien	
<400> 75	
tacttgacct tctttgaaaaa gcattccaa aatgtcttat tttagataga ttaacattaa	60
ccaaacataat tttttttaga tcgagtgc aataatttct aagtcagcct ctagtcgtgg	120
ttcatctctt tcacctgtcat ttatTTGgt gtttgtctga agaaaggaaa gagggaaagca	180
aatacgaatt gtactatTTG taccaaatct ttgggattca ttggcaaata atttcagtgt	240
ggtgttattat t	251
<210> 76	
<211> 251	
<212> DNA	
<213> Homo sapien	
<400> 76	
tatTTAATAA tacaccacac tgaaattatt tgccaatgaa tcccaaagat ttggTacaAA	60
tagtacaatt cgtatTTGct ttccTCTTTC ctttcttcag acaaacacca aataaaatgc	120
aggTgaaaga gatgaaccac gactagaggc tgacttagaa atttatgctg actcgatcta	180
aaaaaaatTA tggTggtaa tgTTaatcta tctaaaatag agcattttgg gaatgcttt	240
caaagaaggt c	251
<210> 77	
<211> 351	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	

<222> (1)...(351)
 <223> n = A,T,C or G

<400> 77

actcaccgtg	ctgtgtgctg	tgtgcctgct	gcctggcagc	ctggccctgc	cgctgctcag	60
gaggcgggag	gcatgagtga	gctacagtgg	gaacaggctc	aggactatct	caagagannn	120
tatctctatg	actcagaaac	aaaaaatgcc	aacagttag	aagccaaact	caaggagatg	180
caaaaattct	ttggcctacc	tataactgga	atgttaaact	cccgcgcat	agaaaataatg	240
cagaagccca	gatgtggagt	gccagatgtt	gcagaataact	cactattcc	aaatagccca	300
aaatgactt	ccaaagtgg	cacctacagg	atcgtatcat	atactcgaga	c	351

<210> 78

<211> 1574

<212> DNA

<213> Homo sapien

<400> 78

gccctgggg	cggaggggag	gggcccacca	cggcattatt	tccgcgagcg	ccggcactgc	60
ccgctccgag	cccggtctg	tcgggtcccg	agccaacttt	cctgcgtcca	tgcagccccg	120
ccggcaacgg	ctgcccgtc	cctggcccg	gcccaggggc	ccgcgcucca	ccgccccgct	180
gctcgcgtg	ctgctgttc	tcgccccgt	ggcggcgccc	gcggggtccg	gggaccccga	240
cgacctggg	caggctcagg	atgctgggt	cccgcgccagg	ctccctgcagc	aggcgccgcg	300
cgccgcgtt	cacttctca	acttccggc	cggtctgccc	agcgcgtgc	gagtgctggc	360
cgaggtgcag	gaggggcccg	cgtggattaa	tccaaaagag	ggatgtaaag	ttcacgtgg	420
cttcagcaca	gagcgtctaca	accagagtc	tttacttcag	gaaggtgagg	gacgtttggg	480
gaaatgttct	gctcgagtgt	tttcaagaa	tcaaaaaccc	agaccaacta	tcaatgtaac	540
ttgtacacgg	ctccatcgaga	aaaagaaaag	acaacaagag	gattacctgc	tttacaagca	600
aatgaagcaa	ctgaaaaacc	ccttggaaat	agttagcata	cctgataatc	atggacatat	660
tgatccctct	ctgagactca	tctgggattt	ggcttcctt	ggaagctctt	acgtatgtg	720
gaaaaatgaca	acacaggtgt	cacactacta	cttggcacag	ctcactagt	tgaggcagt	780
gaaaactaat	gtgtatacaa	ttgattttga	ttatactgtt	ctacttcatg	aattatcaac	840
acagggaaata	attccctgtc	gcattcaatt	ggctctggta	cctggcaaac	ctcttaaagt	900
gaagtaccac	tgtcaagagc	tacagacacc	agaagaagcc	tccggaactg	aagaaggatc	960
agctgttagta	ccaaacagago	tttagtaattt	ctaaaaagaa	aaaatgatct	ttttccgact	1020
tctaaacaag	tgactatact	agcataaaatc	attcttctag	taaaacagct	aaggatata	1080
cattctaata	atttggaaa	acctatgatt	acaagtaaaa	actcagaaat	gcaaagatgt	1140
tggtttttg	tttctcagtc	tgcttagct	tttaactctg	gaagcgcatt	cacactgaac	1200
tctgctcagt	gctaaacagt	caccaggcagg	ttcctcaggg	tttcagccct	aaaatgtaaa	1260
acctggataa	tcagtgtatg	ttgcaccaga	atcagcattt	tttttttaac	tgcaaaaaat	1320
gatggctct	tctctgaatt	tatatttctc	attcttttga	acatactata	gctaataat	1380
tttatgttgc	taaatttgctt	ctatctagca	tgttaaacaa	agataatata	ctttcgatga	1440
aagtaaatta	tagaaaaaaa	attaaactgtt	ttaaaaagaa	cttgattatg	ttttatgatt	1500
tcaggcaagt	attcattttt	aacttgcac	ctactttaa	ataaaatgttt	acatttctaa	1560
aaaaaaaaaa	aaaaaa					1574

<210> 79

<211> 401

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(401)

<223> n = A,T,C or G

```

<400> 79
catactgtga attgttcttg actccctttc ttgacattca gttttcanaa tttccatctt      60
tcttcggaa ctaatgtgct gttctttga ctgcctgctg ggccagcatc cgattgccag      120
ccagaaaacgt cacactgccc aagatggcca ggtacttcaa ggtctggAAC atgttgagct      180
gagtccagta gacatacatg agtcccagca tagcagcatg tcccaggta aatataatcg      240
tgcttagggc aaaagtgaag ttggagacat tggcaccaat ccggatccac tagttctaga      300
gcggccgcca ccgcgggtgga gctccagct ttgttccctt tagtgagggt taattgcgcg      360
cttggcgtaa tcatggncat agctgtttcc tgtgtgaaat t      401

<210> 80
<211> 301
<212> DNA
<213> Homo sapien

<400> 80
aaaaatgaaa catctatTTT agcagcaaga ggctgtgagg gatggggtag aaaaggcatc      60
ctgagagagt tctagaccga cccaggtcct gtggcacact atacgggtca ggaggggtgg      120
aagacaggcc taagctctag gacggtaat ctccgggcta tttgtggatt tgtagaaac      180
agacattctt ttggcctttt cctggcactg gtgttgcgg caggtggca gaagtgagcc      240
accagtcaCT gtcagtcAT tgccaccaca gatcttcAGC agaatcttcc ggtaatcccc      300
t      301

<210> 81
<211> 301
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G

<400> 81
tagccaggTT gctcaagcta attttatTTt ttcccaacag gatccatttG gaaaatatca      60
agcTTtaga atgtggcagc aagagaaaAGC ggactacgca ggaacggggA gtttgggaga      120
agctctcCTG gtgttgactt agggatgaaAG gctccaggct gctGCCAGAA atggagtCAC      180
cagcagaaga actgnTTCT ctgataagGA tggccacca ttttcaagCT gttcgTTAAA      240
gttacacagg tcTTTCTTGC agcagtaagt accgttagCT cattttccCT caagcgggTT      300
t      301

<210> 82
<211> 201
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(201)
<223> n = A,T,C or G

<400> 82
tcaacagaca aaaaaagTTT attgaataca aaactcaaAG gcatcaacAG tcctggggCC      60
aagagatCCA tggcaggaAG tcaagagtTC tgcttcaggG tcggTCTGGG cagccCTggA      120
agaagtCATT gcacatgaca gtgtatgagTG ccaggAAAAC agcataCTCC tggAAAGTCC      180
acctgctggN cactgnTTCA t      201

```

<210> 83
<211> 251
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(251)
<223> n = A,T,C or G

<400> 83
gttaggagca tactgtgcc atttattata gaatgcagtt aaaaaaaaaaaa ttttgaggtt 60
agcctctcca gttaaaaagc acttaacaag aaacacttgg acagcgatgc aatggtctct 120
cccaaaccgg ctcccctcta ccaagtacgg taaacagggt ttgagaacgt tcaatcaatt 180
tcttgatatg aacaatcaaa gcatttaatg caaacatatt tgcttctcaa anaataaaac 240
catttccaa a 251

<210> 84
<211> 301
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G

<400> 84
agtttataat gtttactat gattttagggc tttttttca aagaacaaaa attataagca 60
taaaaaactca ggtatcagaa agactcaaaa ggctgtttt cactttgtc agattttgtt 120
tccaggcatt aagtgtgtca tacagttgtt gccactgctg ttttccaaat gtccgatgtg 180
tgctatgact gacaactact tttctctggg tctgatcaat tttgcagtt accatttttag 240
ttcttacggc gtcnataaca aatgcttcaa catcatcagc tccaatctga agtcttgctg 300
c 301

<210> 85
<211> 201
<212> DNA
<213> Homo sapien

<400> 85
tattttgtta tgtaacattt attgacatct acccactgca agtataagatg aataagacac 60
agtcacacca taaaggagtt tatccttaaa aggagtggaaa gacattcaaa aaccaactgc 120
aataaaaaag ggtgacataa ttgctaaatg gagtgagga acagtgctta tcaattcttg 180
attgggccac aatgatatac c 201

<210> 86
<211> 301
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(301)

<223> n = A,T,C or G

<400> 86

tttataaaaat attttattta cagtagagct ttacaaaaat agtcttaaat taatacaat	60
ccctttgca atataactta tatgactatc ttctaaaaa cgtgacattc gattataaca	120
cataaaactac atttatagtt gttaagtac cttgttagtat aaatatgtt tcatctttt	180
tttgcataaa ggtacatacc aataacaatg aacaatggac aacaaatctt atttgttat	240
tcttccaatg taaaattcat ctctggccaa aacaaaatta accaaagaaa agtaaaacaa	300
t	301

<210> 87

<211> 351

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(351)

<223> n = A,T,C or G

<400> 87

aaaaaaagatt taagatcata aataggcat tgggtcaca acacatttc gaatcttaaa	60
aaaacaaaaca ttttggctt ctaagaaaaa gactttaaa aaaaatcaat tccctcatca	120
ctgaaaaggac ttgtacattt ttaaacttcc agtctcctaa ggcacagtat ttaatcagaa	180
tgccaatatt accaccctgc tggcangaa ataaagaagc aagggattaa cactaaaaaa	240
aacngccaaa ttcttgaacc aaatcattgg cattttaaa aagggataaa aaaacnggnt	300
aaggggggga gcatttttaag taaagaangg ccaagggtgg tatgccnnga c	351

<210> 88

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(301)

<223> n = A,T,C or G

<400> 88

gttttaggtc ttaccaatt tgattggttt atcaacaggg catgagggtt aaatataatct	60
ttgagggaaag gtaaagtcaa atttgacttc ataggcatc ggcgtcctca ctcctgtgca	120
ttttctggtg gaagcacaca gttaattaac tcaagtgtgg cgntagcgat gcttttcat	180
ggngtcattt atccacttgg tgaacttgca cacttgaatg naaactcctg ggtcattggg	240
ntggccgcaa gggaaaggc cccaaagacac caaaccttgc agggtacctn tgcacaccaa	300
c	301

<210> 89

<211> 591

<212> DNA

<213> Homo sapien

<400> 89

tttttttttt tttttttatt aatcaaatga ttcaaaacaa ccatcattct gtcaatgccc	60
aagcaccagg ctggtcctct ccccacatgt cacactctcc tcagcctctc ccccaaccct	120
gctctccctc ctccccctgcc ctagccagg gacagagtct aggaggagcc tggggcagag	180

ctggaggcg	gaagagagca	ctggacagac	agctatggtt	tggattgggg	aagagattag	240
gaagtaggtt	cattaaagacc	cttttttagt	accagatatac	cagccatatt	cccagctcca	300
ttattcaaat	catttccccat	agcccagctc	ctctctgttc	tccccctact	accaattctt	360
tggctctac	acaatttta	tccctcaaat	attcatccct	ggcccaacca	gtccccctgag	420
cctccctctg	gtggagactc	ctccacccat	gagctccccca	gagcatccaa	gacagagtgc	480
acagagacct	ggggaaaggaa	gctgaacttt	gcagagatgt	ggacagggtgc	aggctagggt	540
acagggttgtt	ggttagaggag	acaagttta	tttccaggcc	cacagtctct	c	591

<210> 90
<211> 1978
<212> DNA
<213> *Homo sapien*

<400>	90					
ttttttttt	tttttatca	aatgaatact	ttatttagaga	cataacacgt	ataaaataaa	60
tttctttca	tcatggagtt	accagatttt	aaaaccaacc	aacactttct	cattttaca	120
gctaagacat	gttaaattct	taaatgcac	aattttgtt	caactgctt	gtcattcaac	180
tcacaagtct	agaatgtat	taagctacaa	atctaagtat	tcacagatgt	gtcttaggct	240
tggtttcaa	caatctagaa	gcaatctgtt	tacaaaagtg	ccaccaaagg	attttaaaga	300
aaccaattt	atgccaccaa	acataaggct	gctatacctg	gaaaacaaaa	aatctcacac	360
ctaaattctt	gcagagtaaa	cgattccaa	tagaatgtac	tgtatatcca	tatggcacat	420
ttatgactt	gtaatatgtt	attcataata	caggtttagg	tgtgtggtat	ggagcttagga	480
aaaccaaagt	agtaggatat	tatagaaaaag	atctgtatgtt	aagtataaag	tcatatgcct	540
gatttcctca	aacctttgt	ttttcctcat	gtctctgtc	tttatatttt	tatcacaac	600
caagatctaa	cagggttctt	tctagaggt	tattagataa	gtaacacttg	atcattaagc	660
acggatcatg	ccactcattc	atgggtgtc	tatgttccat	gaactctaatt	agcccaactt	720
atacatggca	ctccaaagggg	atgcttcagc	cagaaagtaa	agggctgaaa	aagttagaaca	780
atacaaaggc	cctcgtgtgg	tggaaactgt	ggccctcactc	ttacttgtcc	ttccattcaa	840
aacagtttgg	cacctttcca	tgacgaggat	ctctacaggt	aggttaaaat	actttctgt	900
gctattcagc	cagaaaatagt	ttttgtgtcg	gatatgattt	taaaacagat	tttgcgtgtc	960
accagtgc当地	aaacattaca	gatgtctggg	ctaataaaaa	aacacataag	aatctacaac	1020
tttatattta	atactctatt	caaatttaac	tcaaagtaat	gaaaataat	tagaagtaaa	1080
aacttaattt	ttctgagac	tctatttgg	aaagcttac	atatccacac	acaaatatgg	1140
gtatattcat	gcacagggca	aacaactgt	ttctgaagca	taaataaaact	caaagtaaga	1200
catcgttagc	tagataccag	ttccagtatt	ggtaatgg	ctctggggat	cccatattaa	1260
gcactctcag	atgaggatct	tgctcagtg	ttagactatc	attagttga	ttaagcaact	1320
gaagtttact	tcataaaattt	cttttccta	tatccaggac	tctgcctgag	aaattttata	1380
cattcctcca	aaggtaagta	ttctccaaag	gtaagtattt	gactattaac	acaaaggca	1440
tgtgattatt	gcataatgac	actaaatatt	atgtggctt	tctgttaggt	ttataagttt	1500
tcaatgtca	gttcaagaaa	atgcagatca	tatataacta	aggtttaca	ccagtggttg	1560
acaaactatg	gccccacaggc	taaaccacgc	ctcccctgt	tttataaaat	aagttttatt	1620
agacataacc	acactcattc	atttctgtat	tgtgtatagc	tgctttcagc	ctataactagc	1680
agaactgaat	agttgtgaca	gagactgtat	ggaccgtgaa	gcataaaat	ttaccatctg	1740
gcccattcta	aaaaaaagtgt	gccaattcct	ggtttacact	aaaatataaga	gtttagtggg	1800
aaggcttattt	gaaatgtgtt	tttttaggg	gctgtatatt	ccaattaaaa	ttaaggttca	1860
ggtgactcag	caaccaaaca	aaagggatac	taattttta	tgaacaat	atttgtat	1920
tatqgacata	aaaggaaact	ttcagaaaga	aaaggaggaa	aataaaagggg	gaaaggga	1978

```
<210> 91  
<211> 895  
<212> DNA  
<213> Homo sapien
```

<400> 91 tttttttttt ttttttcttg tttaaaaaaaa ttgttttcat tttaatgatc tgagtttagta 60

acaaacaaaat	gtacaaaatt	gtctttcaca	tttccataca	ttgtgttatg	gaccaaSatga	120
aaacgcgtgga	ctacaaaatgc	aggtttcttt	atatccttaa	cttcaattat	tgtcaactt	180
aaataaaggt	gatttgctaa	cacatgcatt	tgtgaacaca	gatgcggaaa	attatacatg	240
taagttaatg	cacaaccaag	agtatacact	gttcatttgc	gcagttatgc	gtcaatgcg	300
actgacacag	aagcagttat	cctgggatat	ttcactctat	atgaaaagca	tcttggagaa	360
atagattgaa	atacagttta	aaacaaaaat	tgtattctac	aaataacaata	aaatttgcaa	420
cttgcacatc	tgaagcaaca	tttgagaaag	ctgcttcaat	aaccctgctg	ttatattgg	480
tttataggta	tatctccaaa	gtcatgggtt	gggatatagc	tgctttaaag	aaaataaata	540
tgtatattaa	aaggaaaatc	acactttaaa	aatgtgagga	aagctttgaa	aacagtctta	600
atgcatgagt	ccatctacat	atttcaagt	tttggaaaca	gaaaagaagtt	tagaattttc	660
aaagtaatct	gaaaacttcc	taagccattt	taaaaataaga	ttttttccc	catcttcca	720
atgtttccct	tttgatagtg	taatacagaa	atggcagtt	tctagtgtca	acttaactgt	780
gctaattccat	aagtccattat	acatttatga	cttaagagtt	caaataagtg	gaaattgggt	840
tataatgaaa	atgacaagggg	ggcccccttca	gcagccactc	atctgaacta	gtaat	895

<210> 92

<211> 1692

<212> DNA

<213> Homo sapien

<400> 92

tttttttttt	ttttaactt	tttagcagtgt	ttatttttgt	taaaagaaaac	caattgaatt	60
gaaggtcaag	acaccttctg	attgcacaga	ttaaaaacaaga	aagtattact	tatttcaact	120
ttacaaaagca	tcttatttgc	ttaaaaagat	ccatactatt	gataaaggtc	accatgaaca	180
tatatgtaat	aaggagacta	aaatatttcat	tttacatatac	tacaacatgt	atttcatatt	240
tctaattcaac	cacaatcat	ataggaaaat	atttaggtcc	atgaaaagat	ttcaaaacat	300
taaaaaatta	aagtttgaa	acaaatcaca	tgtaaaagct	cattaaataa	taacattgac	360
aaataaaatag	ttaatcagct	ttacttatta	gctgctgcca	tgcatttctg	gcattccatt	420
ccaaagcgagg	gtcagcatgc	agggtataat	ttcatactat	gkgaccgtaa	agagctacag	480
ggcttatttt	tgaagtggaa	tgtcacaggg	tcttcattt	tcttcaag	gaagatca	540
catggctgct	aaactgttcc	catgaagagt	acccaaaaag	cacctttctg	aaatgttact	600
gtgaagattc	atgacaacat	attttttta	acctgttttgc	aaggagttt	gtttaggaga	660
ggggatgggc	cagtagatgg	agggtatctg	agaagccctt	ttctgttttta	aatataatg	720
attcactgat	gtttatagta	tcaacagtct	ttttaagaaca	atgaggaatt	aaaactacag	780
gatacgtgga	atttaaatgc	aaattgcatt	catggatata	cctacatctt	gaaaaacttg	840
aaaaggaaaa	actattccca	aagaaggcgc	tgatacttaa	gacagctgc	tgggtttgat	900
caaagcagaa	agcatatact	ttcaagttag	aaaacagcag	tggcaggctt	gagtcttcca	960
agcaatcaaa	tctgtaaagc	agatggttac	tagtaagtct	agttatggga	gtctgagttc	1020
taactcatgc	tgtgcttgc	ggatttgctg	gctctttcc	gctctctgt	atgctggact	1080
ggcttgcag	gtgacatgct	ctcaaagtttgc	tgactggact	cggtgtgt	ccgggtgtac	1140
ctcttgcact	tgcaggcgt	gactactgt	attttgtagg	tgcgtgtgt	gccatcttgg	1200
cactgcagct	ggatttcttgc	ggtacgggtt	ttgtcatttgc	cacaccgcca	ctcctggag	1260
ctcctcctgc	tccagttactt	tgttccatag	cctccctccaa	tccagttagg	gagcaactggc	1320
agggggcaagc	actcgccagc	acacaccagc	tccttcagag	ggctgtatgt	ggtgcactgg	1380
ccatcagaga	tgtatgggt	ggaacgcagt	tcccgccaaac	ccacttgaac	ccgagtgttc	1440
cgatccagtc	cagtgttact	gaaatgcctg	cctccatttc	tggcttgatt	caacgtgctg	1500
ttgctgctgg	ggtgtgctgg	aacaggttta	accacatgtg	aataaaggat	ttctgtggca	1560
tcatttttaa	aagccaaaca	gcttttccatt	aggatgcatg	caaggggaag	gagatagaaa	1620
tgaatggcag	gaggaagcat	ggtgagtaga	ggatttgctt	gactgaagag	ctggtaatt	1680
cttttgcctc	tg					1692

<210> 93

<211> 251

<212> DNA

<213> Homo sapien

<400> 93
cccacccat ccaaataattt gacaccaaca cagaaaagct agcaatggat tcccttctac 60
tttgttaaat aaataagtta aatatttaaa tgccctgtgc tctgtgatgg caacagaagg 120
accaacaggc cacatcctga taaaaggtaa gaggggggtg gatcagcaaa aagacagtgc 180
tgtggctga ggggacctgg ttcttgatgt ttccccctca agactcttcc cctacaaata 240
actttcatat g 251

<210> 94
<211> 735
<212> DNA
<213> Homo sapien

<400> 94
tttttttttt tttttccact ttcagttta ttctggac taaatttggg tcagagctgc 60
agagaaggga tggccctga gtttgaggat gaaagtgc cagggagatt gagacgcaac 120
ccccgcctg gacagtttg gaaattgttc ccagggttca actagagaga cacggtcagc 180
ccaatgtggg ggaagcagac cctgagttca ggagacatgg ggtcaggggc tggagagatg 240
aacattctca acatctctgg gaaggaatga ggtctgaaa ggagtgtcag ggctgtccct 300
gcagcagggtg gggatgccgg tgcgtcgat cctggatga ctcaggagtt ggcctggatg 360
gttccttgg tccacttggt gaaacttgcag aggttcgtgt agacacccgg tctgttggc 420
cgggcacaag gtaatctcc ccaggacacg agtccctgca gggagccatt gcagaccaca 480
ggccccccag aatcacccctg gcaggagtct ctacctgct tgcacccggc gcagaacatg 540
gtgtcatcta tctgtctcgg gtaagcatcc tcgcacctt tctgacttag cacgtgtata 600
ttcaaggact ggaggacctt agggaaagtgc acttggggc tcttgggttgc ccccaagcca 660
gacaccaaggc actttgtccc agcagaggga caatgagagg agacgttgc ggtctgaca 720
tcttagtgg gacga 735

<210> 95
<211> 578
<212> DNA
<213> Homo sapien

<400> 95
cttgccttct ctaggcttt gaagcatttt tgtctgtgc ccctgatctt caggtcacca 60
ccatgaagtt ctagcagtc ctggtaactt tggagtttc catctttctg gtctctgccc 120
agaatccgac aacagctgct ccagctgaca cgtatccagc tactggctt gctgatgtg 180
aagccctga tgctgaaacc actgctgtg caaccactgc gaccactgt gtcctacca 240
ctgcaaccac cgctgtttct accactgctc gtaaagacat tccagttta cccaaatggg 300
ttggggatct cccgaatggt agagtgtgc cctggatgg aatcagctt agtcttctgc 360
aattggtcac aactattcat gttccctgtt attcatcca actacttacc ttgcctacga 420
tatccccctt atctctaattt cttaaaata aaaaataact atgagcaaca 480
aaaaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 540
aaaaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 578

<210> 96
<211> 594
<212> DNA
<213> Homo sapien

<400> 96
atggcaaaga atggacttgt aatttgcattt ctgggtatca ctttactcct ggaccagacc 60
accagccaca catccagatt aaaagccagg aagcacagca aacgtcgatg gagagacaag 120
gatggagatc tgaagactca aattgaaaag ctctggacag aagtcaatgc cttgaaggaa 180
attcaagccc tgcagacagt ctgtctccga ggcactaaag ttcacaagaa atgctacctt 240

gcttcagaag	gtttgaagca	tttccatgag	gccaatgaag	actgcattc	caaaggagga	300
atcctgtta	tccccaggaa	ctccgacgaa	atcaacgccc	tccaagacta	tggtaaaagg	360
agcctgccag	gtgtcaatga	ctttggctg	ggcatcaatg	acatggtcac	ggaaggcaag	420
tttgttacg	tcaacggaat	cgcatactcc	ttcctaact	gggaccgtgc	acagcctaac	480
ggtggcaagc	gagaaaaactg	tgtcctgttc	tcccaatcg	ctcagggcaa	gtggagtgtat	540
gaggcctgtc	gcagcagcaa	gagatacata	tgcgagttca	ccatccctca	atag	594

<210> 97
<211> 3101
<212> DNA
<213> Homo sapien

<400> 97						
tgttggggcc	tcagcctccc	aagtagctgg	gactacaggt	gcctgccacc	acgcccagct	60
aatttttgt	atattttta	gtagagacgg	ggtttcacccg	tggtctcaat	ctcctgacct	120
cgtatctgc	caggcttggc	ctcccaaagt	gtattcttt	tttatttatta	tttatttattt	180
tgagatggag	tctgtctctg	tcgcccaggc	tggagtgcag	tggtgcgatc	tctgctcact	240
gcaagctccg	cctcctgggt	tcatgcccatt	ctcctgcctc	agcctcccgaa	gtagctggaa	300
ctacagcccc	ctggcaccac	acccggctaa	ttttttgtat	tttttagtata	gacagggttt	360
caccatgtta	gccagggtgg	tctctatctt	ctgacctcgat	gatccgcctg	cctcagtctc	420
tcaaagtgt	gggattacag	gcgtgagcca	ccgcgaccag	ccaactatgg	ctgtttatTT	480
ttaaatatat	tttaaagaaa	caatttagatt	tgtttttttt	ctcattcttt	tacttctact	540
cttcatgtat	gtataattat	atttgtgttt	tctattacct	tttctccttt	tactgtattt	600
gactataata	atttgtctca	ctaatttctg	ttcactaata	ttatcagctt	agataataact	660
ttaattttta	acttatatat	tgagtattaa	attgatcagt	tttattttgtat	attatctatc	720
ttccgccttgg	ctgaatataa	cttcttaaagc	ttataacttc	ttgttctttc	catgttattt	780
ttttcttttt	tttaatgtat	tgaatttctt	ctgacactca	ttctagtaac	ttttttctcg	840
gtgtgcacacg	taagttataa	tttgccttc	agatttgaga	tctgcataaa	gtttgaggct	900
ttatTTTTT	tttttatTTT	ctttatggca	agtcggacaa	cctgcatgga	tttggcatca	960
atgtagtcac	ccatatctaa	gagcagcact	tgcttcctttag	catgatgagt	tgtttctggaa	1020
ttgtttcttt	attttactta	tattccttgt	agattcttat	attttccctt	caactctatt	1080
cagcatttta	ggaattctta	ggactttctg	agaatttttag	ctttctgtat	taaatgtttt	1140
taatgagttat	tgcattttct	caaaaagcac	aaatatcaat	agtgtacaca	tgaggaaaac	1200
tatatatata	ttctgttgc	gatgacagca	tctcataaaca	aaatcctagt	tacttcattt	1260
aaaagacagc	tctccctcaa	tatactatga	ggtaacaaaa	attttagtgc	tgtatTTTT	1320
ttaatattag	aaaactcatc	ttacattgtg	cacaaaatttc	tgaagtgata	ataacttcact	1380
gtttttctat	agaagtaact	taatattygc	aaaattactt	atttgaattt	aggTTTGGC	1440
tttcatcata	tacttcctca	ttaacatttc	cctcaatcca	taaatgcaat	ctcagtttga	1500
atcttcatt	taacccagaa	gttaattttt	aaaacctttaa	taaaatttga	atgttagctag	1560
atattatttgc	ttggttacat	attagtcaat	aatttatatt	acttacaatg	atcagaaaaat	1620
atgatctgaa	tttctgtgt	cataaattca	ataaacgtatt	ttaggcctaa	acctttccat	1680
ttcaaatcct	ttgggtcttgt	aattgaaaat	aatcattatc	ttttgttttc	tggccaaaaaa	1740
tgctgccccat	ttattttctat	ccctaattag	tcaaaacttc	taataaatgt	attnaacgtt	1800
aatgatgttt	atttgttgt	tgtataactaa	aaccattagt	ttctataatt	taaatgtcact	1860
ctaatatgag	tgaaaatgtg	tcagaggctg	gggaagaatg	tggatggaga	aaggaaaggt	1920
gttgatcaaa	aagtacccaa	gtttcaggtt	cacaggaggc	atgagattga	tctagtgc当地	1980
aaaatgtat	gtataataaa	taataatgca	ctgtatattt	tgaatttgct	aaaatgtat	2040
ttaaaaattgt	tttacataat	attttacata	tttataaaac	acatgcaata	tgttggtaca	2100
tgtatagaat	gtgcaacgat	caagtcaagg	tatctgttgt	atccaccact	ttgagcattt	2160
atcgattcta	tatgtcagga	acatttcaag	ttatctgttc	tagcaaggaa	atataaaaata	2220
cattatagtt	aactatggcc	tatctacagt	gcaactaaac	actagatttt	attcctttcc	2280
aactgtgggt	ttgttattcat	ttaccaccc	cttttcattc	cctttctcac	ccacacactg	2340
tgccggccct	caggcatata	ctattctact	gtctgtctct	gtaaggatata	tcatttttagc	2400
ttccacatata	gagagaatgc	atgcaagtt	tttctttcca	tgtctggctt	atttacttta	2460
acaaaatgac	ctccgcttcc	atccatgtt	tttatattac	ccaatagtgt	tcataaaat	2520

atatacacac atatatacc aattgcattt gtccaattat tcattgacgg aaactggta	2580
atgttatatac gtgtctattt tgaatagtgc tgcaataaac acgcaagttt ggatataattt	2640
tgaagagttt ttgtgttgc gttccataca aattttaaaga ttgttttgc tatgttttg	2700
aaaatggcgtagtattttc atagagattt cattgaatct gtagattgtt ttgggttaagt	2760
atggtttattt tgatgttattt aatttttca ttccatgaag atgagatgtc ttccatttt	2820
tttgtgtcct ctacattttc ttcatcaaa gttttgttattt atttttaaag tagatgttattt	2880
tcacccataa gatcaagtgtt attccctaaa tattttttttt ttgttagctat tgtagatgaa	2940
atggccctctt cgatttctt ttcaactaat tcattattttt tgtagatgaaa tgtagatggat	3000
ttttttttgt tggtttttaa tcaaaaactg tattaaactt agagttttt gtggagttt	3060
taagtttttc tagatataag atcatgacat ctacaaaaaa a	3101

<210> 98
 <211> 90
 <212> PRT
 <213> Homo sapien

<400> 98			
Met Lys Phe Leu Ala Val Leu Val Leu Gly Val Ser Ile Phe Leu			
1	5	10	15
Val Ser Ala Gln Asn Pro Thr Thr Ala Ala Pro Ala Asp Thr Tyr Pro			
20	25	30	
Ala Thr Gly Pro Ala Asp Asp Glu Ala Pro Asp Ala Glu Thr Thr Ala			
35	40	45	
Ala Ala Thr Thr Ala Thr Thr Ala Ala Pro Thr Thr Ala Thr Thr Ala			
50	55	60	
Ala Ser Thr Thr Ala Arg Lys Asp Ile Pro Val Leu Pro Lys Trp Val			
65	70	75	80
Gly Asp Leu Pro Asn Gly Arg Val Cys Pro			
85	90		

<210> 99
 <211> 197
 <212> PRT
 <213> Homo sapien

<400> 99			
Met Ala Lys Asn Gly Leu Val Ile Cys Ile Leu Val Ile Thr Leu Leu			
1	5	10	15
Leu Asp Gln Thr Thr Ser His Thr Ser Arg Leu Lys Ala Arg Lys His			
20	25	30	
Ser Lys Arg Arg Val Arg Asp Lys Asp Gly Asp Leu Lys Thr Gln Ile			
35	40	45	
Glu Lys Leu Trp Thr Glu Val Asn Ala Leu Lys Glu Ile Gln Ala Leu			
50	55	60	
Gln Thr Val Cys Leu Arg Gly Thr Lys Val His Lys Lys Cys Tyr Leu			
65	70	75	80
Ala Ser Glu Gly Leu Lys His Phe His Glu Ala Asn Glu Asp Cys Ile			
85	90	95	
Ser Lys Gly Gly Ile Leu Val Ile Pro Arg Asn Ser Asp Glu Ile Asn			
100	105	110	
Ala Leu Gln Asp Tyr Gly Lys Arg Ser Leu Pro Gly Val Asn Asp Phe			
115	120	125	
Trp Leu Gly Ile Asn Asp Met Val Thr Glu Gly Lys Phe Val Asp Val			
130	135	140	
Asn Gly Ile Ala Ile Ser Phe Leu Asn Trp Asp Arg Ala Gln Pro Asn			

145	150	155	160
Gly	Gly	Lys	Arg
Asn	Cys	Val	Leu
Phe	Ser	Gln	Ser
Ala	Gln	Gly	
165		170	175
Lys	Trp	Ser	Asp
Glu	Ala	Cys	Arg
Ser	Ser	Lys	Arg
Tyr	Ile	Cys	Glu
180		185	190
Phe	Thr	Ile	Pro
		Gln	
195			

<210> 100
<211> 3410
<212> DNA
<213> Homo sapien

<400> 100

gggaaccagc	ctgcacgcgc	tggctccggg	tgacagccgc	gcccctcgcc	caggatctga	60
gtgtatgagac	gtgtcccccac	tgaggtgcc	cacagcagca	ggtgttggc	atgggctgag	120
aagctggacc	ggcacccaaag	ggctggcaga	aatgggcgc	tggctgattc	ctaggcagtt	180
ggcggcagca	aggaggagag	gcccgcagtt	ctggagcaga	gccgagacga	agcagttctg	240
gagtgcctga	acggccccct	gagccctacc	cgccctggccc	actatggtc	agaggctgtg	300
ggtgagccgc	ctgctgcggc	accggaaagc	ccagctttg	ctggtaaacc	tgctaaccctt	360
tggcctggag	gtgtgtttgg	ccgcaggcat	cacctatgt	ccgcctctgc	tgcttggaa	420
gggggttagag	gagaagttca	tgaccatgg	gctgggcatt	ggtccagtg	tgggcttgg	480
ctgtgtcccg	ctccctaggct	cagccagtga	ccactggcg	ggacgctatg	gccggccggc	540
gccccttcattc	tgggcactgt	ccttggccat	cctgctgagc	ctttttctca	tcccaagggc	600
cggctggcta	gcagggctgc	tgtggccga	tccaggccc	ctggagctgg	cactgctcat	660
cctggggctg	gggctgctgg	acttctgtgg	ccaggtgtgc	ttcaactccac	tggaggccct	720
gctctctgac	ctcttccggg	acccggacca	ctgtcgccag	gcctactctg	tctatgcctt	780
catgatcagt	cttgggggct	gcctgggcta	cctccctgc	gccattgtact	gggacaccag	840
tggcctggcc	cccttacctgg	gcacccagga	ggagtgcctc	tttggcctgc	tcaccctcat	900
cttccttacc	tgcgttagcag	ccacactgt	ggtggctgag	gaggcagcgc	tggggccac	960
cgagccagca	gaagggctgt	cgccccccctc	cttgcggccc	cactgctgtc	catgggggc	1020
ccgcttggct	ttccggaaacc	tggggccct	gcttccccgg	ctgcaccagc	tgtgtgcgg	1080
catgccccgc	accctgcgccc	ggctttctgt	ggctgagctg	tgcagctgga	tggcactcat	1140
gaccttcacg	ctgttttaca	cggatttcgt	gggcgagggg	ctgttaccagg	gcgtgcccag	1200
agctgagccg	ggcacccgagg	cccgagaca	ctatgtgaa	ggcggtcgga	tgggagccct	1260
ggggctgttc	ctgcagtgcg	ccatctccct	ggtcttctct	ctggtcatgg	accggctgg	1320
gcagcgattc	ggcactcgag	cagtctattt	ggccagtg	gcagctttcc	ctgtggctgc	1380
cgggtccaca	tgcctgtccc	acagtgtggc	cgtgttgaca	gcttcagccg	cccttaccgg	1440
gttcacccctc	tcagccctgc	agatcctgca	ctacacactg	gcctccctct	accaccggga	1500
gaagcaggtg	ttcctgcca	aataccgagg	ggacactgga	ggtgctagca	gtgaggacag	1560
cctgtatgacc	agcttccctgc	caggccctaa	gcctggagct	cccttccctta	atggacacgt	1620
gggtgttgg	ggcagtggcc	tgctccacc	tccaccggc	ctctgcgggg	cctctgcctg	1680
tgtatgttcc	gtacgtgtgg	tgggtggtga	gccaccggag	gccagggtgg	ttccggggccg	1740
gggcatctgc	ctggacctcg	ccatcttgg	tagtgcctc	ctgtgttccc	aggtggccccc	1800
atccctgttt	atgggctcca	ttgtccagct	cagccagtct	gtcaactgcct	atatgggtgc	1860
tgccgcaggc	ctgggtctgg	tcgcattta	ctttgttaca	caggtagtat	ttgacaagag	1920
cgacttggcc	aaataactcag	cgtaaaaaac	ttccagcaca	ttgggggtgg	gggcctgcct	1980
cactgggtcc	cagctccccg	ctccctgttag	ccccatgggg	ctggccggct	ggccgcccagt	2040
ttctgtgtct	gccaaagtaa	tgtggctetc	tgctgccacc	ctgtgtctgt	gagggtgcgt	2100
gctgcacagc	tgggggctgg	ggcgtccctc	tcctctctcc	ccagtcctta	gggcgtgcctg	2160
actggaggcc	ttccaaggggg	gtttcagtct	ggacttatac	agggaggccca	gaagggtctcc	2220
atgcacttgg	atgcggggac	tctgcagggt	gattaccagg	gctcagggtt	aacagctagc	2280
ctccttagtt	agacacacact	agagaagggt	tttgggagc	tgaataaaact	cagtcacactg	2340
tttcccatc	tctaagccccc	ttaacctgca	gcttcgttta	atgtagctct	tgcacatggag	2400
tttcttaggt	gaaacactcc	tccatgggat	ttgaacatat	gacttatttg	tagggaaaga	2460

gtcctgaggg gcaacacaca agaacccagg tccctcagcc cacagcaactg tcttttgc	2520
gatccacccc cctcttacct tttatcagga tgtgcctgt tggccttct gttccatca	2580
cagagacaca ggcattaaa tatttaactt atttattaa caaagttagaa ggaaatccat	2640
tgctagctt tctgttgtgg tgtctaataat ttgggttaggg tggggatcc ccaacaatca	2700
ggtccccctga gatagctggt cattgggctg atcattgcca gaatcttctt ctccctgggt	2760
ctggcccccc aaaatgccta acccaggacc ttggaaattc tactcatccc aaatgataat	2820
tccaaatgct gttacccaag gttaggggtgt tgaaggaagg tagagggtgg ggcttcaggt	2880
ctcaacggct tccttaacca cccctcttct cttggcccag cctggttccc cccacttcca	2940
ctccccctcta ctctctctag gactgggctg atgaaggcac tgcccaaaat ttcccctacc	3000
cccaacttcc ccctacccccc aactttcccc accagctcca caaccctgtt tggagctact	3060
gcaggaccag aagcacaaag tgcggtttcc caaggctttg tccatctcag ccccccagagt	3120
atatctgtgc ttggggaaatc tcacacagaa actcaaggagc acccccctgcc tgagctaagg	3180
gaggtcttat ctctcagggg gggtttaagt gccgttgca ataatgtcgt cttatttatt	3240
tagcgggggtg aatattttat actgtaaagtg agcaatcaga gtataatgtt tatggtgaca	3300
aaattaaagg ctttcttata tgttttaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa	3360
aaaaaaaaara aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaataa aaaaaaaaaaaa	3410

<210> 101
 <211> 553
 <212> PRT
 <213> Homo sapien

<400> 101	
Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys Ala	
1 5 10 15	
Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu	
20 25 30	
Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val	
35 40 45	
Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly Pro Val Leu Gly	
50 55 60	
Leu Val Cys Val Pro Leu Leu Gly Ser Ala Ser Asp His Trp Arg Gly	
65 70 75 80	
Arg Tyr Gly Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Ile	
85 90 95	
Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu	
100 105 110	
Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly	
115 120 125	
Val Gly Leu Leu Asp Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu	
130 135 140	
Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala	
145 150 155 160	
Tyr Ser Val Tyr Ala Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr	
165 170 175	
Leu Leu Pro Ala Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu	
180 185 190	
Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu	
195 200 205	
Thr Cys Val Ala Ala Thr Leu Leu Val Ala Glu Ala Ala Leu Gly	
210 215 220	
Pro Thr Glu Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His	
225 230 235 240	
Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu	
245 250 255	

Leu Pro Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg
 260 265 270
 Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe
 275 280 285
 Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val
 290 295 300
 Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly
 305 310 315 320
 Val Arg Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu
 325 330 335
 Val Phe Ser Leu Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg
 340 345 350
 Ala Val Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala
 355 360 365
 Thr Cys Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu
 370 375 380
 Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala
 385 390 395 400
 Ser Leu Tyr His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly
 405 410 415
 Asp Thr Gly Gly Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu
 420 425 430
 Pro Gly Pro Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala
 435 440 445
 Gly Gly Ser Gly Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser
 450 455 460
 Ala Cys Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala
 465 470 475 480
 Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp
 485 490 495
 Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser
 500 505 510
 Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala
 515 520 525
 Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp
 530 535 540
 Lys Ser Asp Leu Ala Lys Tyr Ser Ala
 545 550

<210> 102

<211> 940

<212> DNA

<213> Human

<400> 102

tttactgctt ggcaaagtac cctgagcatc agcagagatc ccgagatgaa atcaggaaac	60
tcctaggggta tgggtcttct attaccttggg aacaccttag cccatgcct tacaccacga	120
tgtgcatcaa ggaatgcctc cgcccttacg caccggtagt aaacatatcc cggttactcg	180
acaaaacccat caccttcca gatggacgt ccttacctgc aggaataact gtgttatca	240
atatttgggc tcttcaccac aacccttatt tctggaaaga ccctcaggtc tttaacccct	300
tgagattctc cagggaaaat tctgaaaaaa tacatcccta tgccctcata ccattctcag	360
ctggattaag gaactgcatt gggcagcatt ttgccataat tgagtgtaaa gtggcagtgg	420
cattaactct gctccgcttc aagctggctc cagaccactc aaggcctccc cagcctgttc	480
gtcaaggtgt cctcaagtcc aagaatgaa tccatgtgtt tgcaaaaaaaaaa gtttgctaatt	540
tttaagtcc ttctgtataag aattaatgag acaattttcc taccaaaaggaa agaacaaggaa	600

gataaaatata atacaaaata tatgtatatg gttgttgac aaattatata acttaggata	660
cttctgactg gtttgacat ccattaacag taatttaat ttcttgctg tatctggta	720
aacccacaaa aacmcctgaa aaaactcaag ctgacttcca ctgcgaaggg aaattattgg	780
tttgtgtAAC tagtggtaga gtggcttca agcatagttt gatcaaaact ccactcagta	840
tctgcattac ttttatyttyt gcaaataatct gcatgatgc tttattytc gttatcttc	900
cccataataa aaaatatctg caaaaaaaaaaaaaaa	940
<210> 103	
<211> 529	
<212> DNA	
<213> Human	
<400> 103	
ttttttttt ttttactga tagatggaat ttattaagct tttcacatgt gatagcacat	60
agtttaatt gcatccaaag tactaacaAA aactctagca atcaaraatg gcagcatgtt	120
attttataac aatcaacacc tggctttt AAAATTGTT tttcataara taatttatac	180
tgaagtaaat ctggccatgc tttaaaaaaa tgcttttaggt cactccaagc ttggcaggtt	240
acatTTGGCA taaaacaataa taaaacaatc acaatttaat aaataacaaa tacacattt	300
taggccataa tcataatacag tataaggAAA aggkgttagt gttgagtaag cagtattag	360
aatagaatac ctggccctct atgcaaaatAT gtctaracac tttgatttac tcagccctgA	420
cattcagttt tcaaaagttagg agacaggTTC tacagtatca ttttacagtt tccaaacacat	480
tggaaaacaag tagaaaatga tgagttgatt tttattaatg cattacatc	529
<210> 104	
<211> 469	
<212> DNA	
<213> Human	
<400> 104	
ccccacacaa tggataaaaaa cacttatagt aaatggggac attcactata atgatctaag	60
aagctacaga ttgtcatagt tggcttcctg ctttacaaaaa ttgctccaga tctgaaatgc	120
cagtttggacc ttgttcttct ataataatttC ctttttttcc cctctttgaa tctctgtata	180
tttgattctt aactaaaattt gttctcttaa atattctgaa tcctggtaat taaaagttt	240
ggtgtatTTT ctttacctcc aaggaaagaa ctactagcta caaaaaataat tttgaaataa	300
gcattttttt ggtataaggt acatattttg gttgaaagaca ccagactgaa gtaaacagct	360
gtgcattccaa ttattatAG ttttgcatagt aacaatatgt aatcaaaactt ctatgtact	420
tgagagtggA acctcctata tcattatTA gcaccgttt tgacagtaa	469
<210> 105	
<211> 744	
<212> DNA	
<213> Human	
<400> 105	
ggcctgggac aggattgagg tatgttgcaG cctccaggGC ctggggctc ctgcataag	60
aatacccccC cccatttgac tggactttt ttggcttggA ttctggagaa cagatttcca	120
ggattgtcag ccagaaggca gacagatgc ggcacctacc aagacactgac ctcaggaagt	180
ggccctgccc tacagcccaG ttgctcagcc aggctgaaAg gccatggggc cccagcaccc	240
ttgcttcagt gccagccccC ggaagGAACc tcacaacagg gatacagcaA ggacactcca	300
gttccccccAG tcctgcccAG gtgctaccctt gaggGAAGC gatggagaca gggcagccag	360
gtttgcagg acctgcataG cggggcccaAG actgcccctc ctcttaagtc atgcAAAGC	420
ctccctgccc agtctgagac agtgcgttcc Aggtgaccac gacctgcgtg gcccctccgg	480
cagttgtcat ggtgggtgtA cccccccccA tccccctgag gagacatggg ctcagtcCcA	540
tgcctgggtcC ccacagccac aaagatggcc atgggtctct agcctgatAt tcgtggcctg	600
gcaggggtca gcacccctga gggcatccaa gccatggta gaggaaagtG ttggcaggct	660

cggcacagcc aaagaagtca ggacccacga gacggggaa gccttcaga gccttcacct	720
tcacagggtc aaacttccag taga	744
<210> 106	
<211> 401	
<212> DNA	
<213> Human	
<400> 106	
acattgttag gtgctgacct agacagagat gaactgaggt ctttgtttg ttttgttcat	60
aatacaaagg tgctaattaa tagtatttca gatactgaa gaatgtttag ggtgcttagaa	120
gaatttgaga agaaatactc ctgtatttag ttgtatcggt tgggttattt tttaaaaaat	180
ttgattttgc attcatattt tccatcttat tcccaattaa aagtatgcag attatttgc	240
caaataatcttct tcagattcag catttgtct ttgcagttct cattttcatc ttcttccatg	300
gttccacaga agctttgttt cttgggcaag cagaaaaatt aaattgtacc tattttgtat	360
atgtgagatg tttaaataaaa ttgtgaaaaa aatgaaataaa a	401
<210> 107	
<211> 1009	
<212> DNA	
<213> Human	
<400> 107	
cgagctatta tggtacggaa ctttttttaa tgaggaattt catgatgatt taggaatttt	60
ctctcttgaa aaaggcttcc cctgtatgaa aaatgatgtg ccagctaaaa ttgtgtgcc	120
tttaaaaaact gaaaatattt taaaattatt tgtcttatatt ctaaatttag ctttggatca	180
aacttttaggc caggaccagc tcattgcgttc tcattcttcc ttttctact ctttctctca	240
tcactcacct ctgtattcat tctgtttttt gggatagaaa aatcataaag agccaacccca	300
tctcagaacg ttgtggattt agagagacac tacatgactc caagtataatg agaaaaaggac	360
agagctctaa ttgataactc tggatgttcaa aagggaaaaga gtatgccc aa ttctctctac	420
atgacatatt gagattttttt ttaatcaact tttaaatgtatg tgatgttctg ttctaaactg	480
ttctgttttta gtgaaggtag atttttataa aacaagcatg gggattctt tctaaggtaa	540
tattaatgtatg aaggggaaaaa agtatacttta acagctttt gttgaagcct gtggtagcmc	600
attatgttta taattgcaca tggcacata atctattatg atccaatgca aatacagctc	660
caaaaatatt aaatgtatat atattttaaa atgcctgagg aaatacattt ttcttaataa	720
actgaagagt ctcaatgtatgg ctattaaat aattattatgc ctccctgttgg gtggctgcaa	780
aacatcacaa agtgaccgggt cttagacact gtgaactgtc gcccgttta gtaaataaaaa	840
ttaatgcatt tctagaggggg gaatatctgc catccagtgg tggaaatgtg gagtaaagaa	900
gctgggtggtc tgcttctgtg ctgtatgcca gcctttgc ttaagttgag aggaggtcaa	960
ctttagctac tggatgttggt ttgagagcca tggcaaaaaa aaaaaaaaaa	1009